From THE INSTITUTE OF ENVIRONMENTAL MEDICINE
Karolinska Institutet, Stockholm, Sweden

INTERACTION OF GENETIC AND ENVIRONMENTAL FACTORS IN CHILDHOOD ASTHMA AND ALLERGY

Anna Gref

Karolinska Institutet

Stockholm 2017

Published by Karolinska Institutet.
Printed by E-print AB 2017

ISBN 978-91-7676-620-0

Interaction of genetic and environmental factors in childhood asthma and allergy

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Anna Gref

Principal Supervisor:
Associate Professor Erik Melén
Karolinska Institutet
Institute of Environmental Medicine
Unit of Environmental Epidemiology

Co-supervisors:
Professor Göran Pershagen
Karolinska Institutet
Institute of Environmental Medicine
Unit of Environmental Epidemiology

Associate Professor Anna Bergström
Karolinska Institutet
Institute of Environmental Medicine
Unit of Environmental Epidemiology

Opponent:
Professor Rosalind J Wright
Icahn School of Medicine at Mount Sinai
Department of Pediatrics
New York, USA

Examination Board:
Professor Kjell Torén
Göteborgs Universitet
Department of Medicine
Division of Public Health and Community Medicine

Professor Karin Broberg
Karolinska Institutet
Institute of Environmental Medicine
Unit of Metals and Health

Professor Eva Rönmark
Umeå Universitet
Department of Public Health and Clinical Medicine
Division of Occupational Medicine

“Our goal should be to understand our differences.”

James D. Watson

ABSTRACT

Asthma and allergy are the most common chronic childhood diseases. Their occurrences are influenced by both genetic and environmental factors, although the underlying mechanisms are not fully understood.

The overall aim of this thesis was to study the role of genetic polymorphisms and epigenetic alterations in interaction with environmental factors related to childhood respiratory and allergic disease. Epidemiological data from the Swedish prospective birth cohort BAMSE and other European and North American cohorts were analyzed, in addition to material from functional studies and bioinformatics resources.

Study I included 14,495 children from five European and one North American birth cohorts. Early mold exposure at home was associated with increased risk of early wheezing and childhood nasal symptoms. Effect modification by the functional *GSTP1* single nucleotide polymorphism (SNP) Ile105Val could not be confirmed. However, stratification by genotype revealed statistically significant associations between mold exposure and childhood nasal symptoms confined to Ile/Ile carriers.

The previously identified asthma-associated gene *NPSR1* was investigated in **Study II.** Differential DNA methylation in the *NPSR1* promoter region was associated with adult severe asthma (CpG site 5) (n=171), and childhood allergic asthma (CpG sites 1 and 4) (n=546) but not with non-allergic asthma. Current and former smoking in adults was associated with increased DNA methylation (CpG sites 1 and 8) as compared with never-smoking. The season of blood sampling was associated with differential methylation at CpG site 5, whereas no such associations were seen for season of birth.

Study III aimed to identify genetic regions of importance in the association between traffic air pollution exposure (NO_2) and childhood asthma, with further functional analyses of identified loci. A genome-wide interaction study (GWIS) in 1,534 children from three European birth cohorts, with look-up of SNPs with interaction $p < 1x10^{-4}$ in two North American cohorts (n = 1,602 and 186, respectively), revealed significant interaction for eight SNPs in four different loci. Two of the identified genes, *ADCY2* and *DLG2*, have previously been associated with lung function and COPD. Functional follow-up showed an inverse association between NO_2 exposure at birth and DNA methylation in one CpG site of *DLG2* in 8-year-olds. Short-term diesel exposure in adults revealed inverse associations with *DLG2* DNA methylation in eight additional CpG sites and positive associations in two CpG sites. One of the SNPs identified in the look-up evaluation, rs686237, was found to strongly influence *B4GALT5* expression in adult lung tissue. Interaction between rs686237 and NO_2 exposure at birth in relation to *B4GALT5* gene expression was detected in blood from 16-year-olds.

Associations between dietary antioxidant intake and development of asthma and allergic disease were investigated in 2,359 children in **Study IV**. An inverse association was identified for new onset sensitization to inhalant allergens between 8 and 16 years of age. The association was stronger among children exposed to low levels of current traffic air pollution (NO_x), with a significant interaction between antioxidant intake and current NO_x.

In conclusion, the results in this thesis imply that genetic variants can influence associations between environmental exposures, such as mold and traffic air pollution, and asthma and allergic disease. Interactions between environmental factors may also influence disease development. Further, there seems to be a link between environmental exposures and epigenetic as well as transcriptomic events, which warrant further investiogations.

LIST OF SCIENTIFIC PAPERS

I. Tischer CG, **Gref A**, Standl M, Bauer M, Bergström A, Brauer M, Carlsten C, Gehring U, Granell R, Henderson J, Kerkhof M, MacNutt M, Melén E, Wickman M, Heinrich J. Glutathione-S-transferase P1, early exposure to mould in relation to respiratory and allergic health outcomes in children from six birth cohorts. A meta-analysis. *Allergy. 2013 Mar;68(3):339-46.*

II. Reinius LE, **Gref A**, Sääf A, Acevedo N, Joerink M, Kupczyk M, D'Amato M, Bergström A, Melén E, Scheynius A, Dahlén SE, Pershagen G, Söderhäll C, Kere J. DNA methylation in the Neuropeptide S Receptor 1 (NPSR1) promoter in relation to asthma and environmental factors. *PLoS One. 2013;8(1):e53877.*

III. **Gref A**, Merid SK, Gruzieva O, Ballereau S, Becker A, Bellander T, Bergström A, Bossé Y, Bottai M, Chan-Yeung M, Fuertes E, Ierodiakonou D, Jiang R, Joly S, Jones M, Kobor MS, Korek M, Kozyrskyj AL, Kumar A, Lemonnier N, MacIntyre E, Ménard C, Nickle D, Obeidat M, Pellet J, Standl M, Sääf A, Söderhäll C, Tiesler CMT, van den Berge M, Vonk JM, Vora H, Xu CJ, Antó JM, Auffray C, Brauer M, Bousquet J, Brunekreef B, Gauderman WJ, Heinrich J, Kere J, Koppelman GH, Postma D, Carlsten C, Pershagen G, Melén E. Genome-wide interaction analysis of air pollution exposure and childhood asthma with functional follow-up. *Am J Respir Crit Care Med. 2016 Nov 30.*

IV. **Gref A**, Rautiainen S, Gruzieva O, Håkansson N, Kull I, Pershagen G, Wickman M, Wolk A, Melén E, Bergström A. Dietary total antioxidant capacity in early school age and subsequent allergic disease. *Clin Exp Allergy. 2017 Feb 21.*

CONTENTS

LIST OF ABBREVIATIONS

2 df	Two-degree-of-freedom
A	Adenine
ADCY2	Adenylate cyclase 2
ALSPAC	Avon Longitudinal Study of Parents and Children
B4GALT5	UDPGal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5
BAMSE	Children, Allergy, Milieu, Stockholm, Epidemiological study
BIOAIR	Longitudinal assessment of clinical course and BIOmarkers in severe chronic AIRway disease
C	Cytosine
CAPPS	Canadian Asthma Primary Prevention Study
CHS	Children's Health Study
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
DEP	Diesel exhaust particles
DLG2	Discs large homolog 2
DNMT	DNA methyltransferase
EMSA	Electrophoretic mobility shift assay
eQTL	Expression quantitative trait locus
EWAS	Epigenome-wide association study
FANTOM5	Functional ANnoTation Of the Mammalian genome
FAP	Filtered air particles
FEV_1	Forced expiratory volume in one second
FFQ	Food frequency questionnaire
FVC	Forced vital capacity
G	Guanine
GINI	German Infant study on the influence of Nutritional Intervention plus environmental and genetic influences on allergy development
GSDMB	Gasdermin B
GST	Glutathione S-transferase
GSTP1	Glutathione S-transferase pi 1
GTEx	Genotype-tissue expression
GWAS	Genome-wide association study
GWIS	Genome-wide interaction study
GxE	Gene-environment interaction
IgE	Immunoglobulin E
IL	Interleukin
Ile	Isoleucine
LD	Linkage disequilibrium
LISA	Lifestyle related factors, Immune System and the development of Allergies in East and West Germany plus the influence of traffic emissions and genetics study
LRT	Likelihood ratio test

LUR	Land use regression
MAF	Minor allele frequency
MAGI1	Membrane-associated guanylate kinase, WW and PDZ domain containing 1
methQTL	Methylation quantitative trait locus
MOCOS	Molybdenum cofactor sulfurase
mRNA	Messenger ribonucleic acid
NO	Nitrogen oxide
NO_2	Nitrogen dioxide
NO_x	Nitrogen oxides
NPSR1	Neuropeptide S Receptor 1
OR	Odds ratio
ORAC	Oxygen radical absorbance capacity
ORMDL3	ORMDL sphingolipid biosynthesis regulator 3
PCR	Polymerase chain reaction
PIAMA	Prevention and Incidence of Asthma and Mite Allergy
$PM_{2.5}$	Particles with an aerodynamic diameter less than 2.5 μm
PM_{10}	Particles with an aerodynamic diameter less than 10 μm
QC	Quality control
ROS	Reactive oxygen species
SAGE	Study of Asthma, Genes, and Environment
SNP	Single nucleotide polymorphism
SPT	Skin prick test
T	Thymine
TAC	Total antioxidant capacity
TAG	Traffic pollution, Asthma and Genetics
TE	Trolox equivalents
TF	Transcription factor
TNF	Tumor necrosis factor
U	Uracil
US	United States
Val	Valine

1 INTRODUCTION

Asthma and allergy in children are complex diseases meaning that both the individual genetic setup and environmental factors contribute to determining disease incidence and persistence. Parental allergic disease is the primary risk factor for disease development in children. Heritability of allergic disease (asthma, rhinitis and eczema) has been estimated to 50-80%[1, 2]. The importance of environmental factors may be reflected by the fact that disease prevalence has increased in the last 50 years[3, 4], most likely due to a rapid change in the exposure to environmental factors. When studying complex diseases like asthma or allergy, one has to consider that some individuals could be more susceptible to certain environmental exposures due to their genetic make-up.

This thesis aimed to explore the complexity of allergic disease by studying interactions between environmental and genetic factors using different epidemiological and experimental approaches.

The methods used in genetic epidemiology have developed in parallel with advances in our understanding of the human genome. Thanks to new genotyping technologies and the organization of large consortia and international collaborations, I have had the possibility to investigate gene-environment interactions (GxE) at a genetic, epigenetic and transcriptional level within this thesis.

2 BACKGROUND

2.1 RESPIRATORY AND ALLERGIC DISEASES

The prevalence of asthma and allergic disease has increased during the second half of the last century, but seems to have reached a plateau in high income countries[3, 4]. Around 30-40% of the population in the world have some type of allergic disease[4]. Asthma is the most common chronic disease in children[5].

2.1.1 Allergic disease

Allergic diseases are disorders manifested by an immune response towards normally harmless substances in the environment, called allergens. Allergic reactions can be viewed as an overreaction of the immune system to these allergens. Allergens are usually proteins that exist at low levels in the environment. These can be inhaled from outdoor air, e.g., pollen or molds, or from indoor air, e.g., allergens from house dust mites or furred pets. A person can also be exposed to allergens through ingestion of foods. Below follows a simplified description of the sensitization process and a subsequent asthmatic allergic reaction upon inhalation of an allergen.

2.1.1.1 The sensitization process

Sensitization refers to production of immunoglobulins (for allergy, mainly immunoglobulin E (IgE)) towards allergens[6, 7]. The first time a person encounters an airborne allergen it will pass through the airway epithelium and be taken up by dendritic cells (antigen presenting cells). The dendritic cell migrates to the regional lymph node, where it presents the allergen to a naïve CD4+ T cell. The naïve T cell is stimulated by the cytokine interleukin 4 (IL4) to develop into a Th2 cell. The Th2 cytokines IL4 and IL13 then stimulate B cells to production of allergen-specific IgE antibodies. These IgE antibodies can bind to mast cells positioned beneath the airway epithelium and to basophils.

The fate of a naïve T cell depends on which cytokines it is exposed to, the properties of the antigen, the dose and the route of exposure as well as the host genotype and co-exposure with other agents that can enhance the sensitization process. Development of a naïve T cell into a Th1 cell will lead to production of interferon gamma (IFNɣ) and inhibition of IgE production.

2.1.1.2 Allergic reaction

A person can be sensitized without showing clinical symptoms after additional exposure to the allergen. However, the probability of clinical symptoms increases with increased levels of IgE antibodies present. A person who does show symptoms after additional inhalation of the allergen is said to be allergic. Re-exposure to the allergen triggers an immediate allergic reaction. Allergen binding to the allergen-specific IgE on a mast cell triggers mast cell activation[7]. The mast cell will then release inflammatory mediators such as histamine, causing increased mucous production, bronchoconstriction, vasodilation and increased vascular permeability with edema in the lung.

This rapid response is followed by a slower reaction with sustained inflammation by activation of transcription factors and recruitment and activation of several cell types from the blood stream, including Th2 cells, eosinophils, and basophils, contributing to a late-phase reaction. This occurs hours after the allergen exposure and is characterized by airway narrowing and mucus hypersecretion. Prolonged or repetitive exposure to the allergen can subsequently evolve into a chronic inflammation state characterized by presence of leucocytes, more mast cells with large amounts of IgE bound and structural changes with increased smooth muscle cells, scar tissue, mucous production and epithelial barrier dysfunction.

Activation of mast cells beneath the airway epithelium will result in allergic asthma, whereas activation of mast cells beneath the nasal epithelium is associated with allergic rhinitis.

2.1.1.3 Allergic sensitization

Investigation of allergic sensitization at the clinic as well as in epidemiological studies can be performed through a skin prick test or by measuring the levels of allergen-specific IgE antibodies in the blood. Up to 40% of the world's population are sensitized to some type of allergen[4]. Among children, the prevalence has been reported to increase by age, and the prevalence in school children has been reported to be up to 50%[4, 8, 9]. In Swedish children, sensitization to birch pollen and furred animals is most common[8, 9], whereas sensitization to molds, house dust mite or cockroaches is common in other countries[4].

2.1.2 Asthma

Asthma is a heterogeneous disease characterized by chronic airway inflammation with structural changes of the airway epithelium such as increased smooth muscle cells, mucus hypersecretion and narrowing of the airways. It is associated with hyper-responsiveness and airflow obstruction resulting in shortness of breath, a wheezing sound at breathing, and cough[10]. In mild asthmatics, the exacerbations can be triggered by respiratory tract infections, physical activity, cold air or exposures such as tobacco smoke or allergens[4]. Severe asthmatics have symptoms that are more continuous despite treatment.

Clinically, asthma is diagnosed based on patient history and symptoms of wheeze, heavy breathing, and cough. Obstructive symptoms are common among young children, especially in combination with airway infections, making it harder to diagnose true asthma in younger ages. Epidemiologists often use standardized questionnaires to assess asthma. Parental-reported doctor's diagnosis of asthma, current symptoms of wheeze, and the use of asthma medication can be used to define previous or current disease.

2.1.2.1 Lung function

Lung function tests can be performed to evaluate a clinical diagnosis, such as asthma or chronic obstructive pulmonary disease (COPD). Spirometry measures the volume of air a person can breathe in or out and the flow. Forced vital capacity (FVC) is the total volume of air exhaled. Forced expiratory volume in one second (FEV_1) is the amount of air exhaled

during the first second and this spirometry parameter is a commonly used measure of lung function. The ratio of FEV_1 on FVC (FEV_1/FVC) gives an indication of airway obstruction that is characteristic of asthma and COPD.

2.1.3 Rhinitis

The clinical characteristics of allergic as well as non-allergic rhinitis include sneezing, itching, and a runny, blocked nose, due to inflammation in the nasal epithelium[11]. In sensitized patients, allergic conjunctivitis commonly co-occurs with allergic rhinitis and is characterized by itchy, red and watery eyes, and sometimes also swelling around the eyes. Allergic rhinitis is most often triggered by allergen exposure, leading to an IgE-mediated inflammation of the nasal mucosa [4].

A clinical history of symptoms and exacerbating factors supported by examination of the nose is used to diagnose rhinitis, and IgE testing for classification of allergic vs. non-allergic rhinitis. In epidemiological studies, questionnaires regarding nasal symptoms without simultaneous common cold (and IgE tests) are used to assess rhinitis.

2.2 OXIDATIVE STRESS AND INFLAMMATION

The exact biological mechanisms and individual susceptibility for asthma and allergic diseases are not fully understood, but antioxidant and inflammatory responses have been suggested to be among the key pathways[12]. The inflammatory response and release of reactive oxygen species (ROS), by bronchial epithelial cells as well as by eosinophils and neutrophils, have been found to be increased in asthmatics as compared with non-asthmatics[13, 14].

Normally, there is a balance in the body between oxidant and antioxidant events. Oxidative stress can be simplified as a shift in the balance towards more oxidative events (**Figure 2.1**). Oxidative stress and acute inflammation are natural defense mechanisms that protect the body from harmful events.

At low levels of oxidative stress, the transcription factor nuclear factor E2-related factor 2 (Nrf2) is activated and induces expression of more than 200 genes, such as the antioxidant enzymes catalase, superoxide dismutase, heme-oxygenase-1 and glutathione s-transferases (GSTs) which can detoxify free radicals[15]. These are endogenous antioxidants. At higher levels of oxidative stress, pro-inflammatory cytokines and chemokines become activated, which can lead to increased airway and systemic inflammation[16].

Diet can contribute with exogenous antioxidants, such as vitamin C, vitamin E, carotenoids and polyphenols, but there are many more antioxidants present in our diet, which can interact with each other in synergistic or antagonistic manners[17]. Dietary antioxidants are suggested to improve endogenous antioxidant defenses and lower oxidative stress levels. Diminished

antioxidant protection or induction of ROS from environmental exposures, such as air pollution, might induce oxidative stress and inflammatory responses.

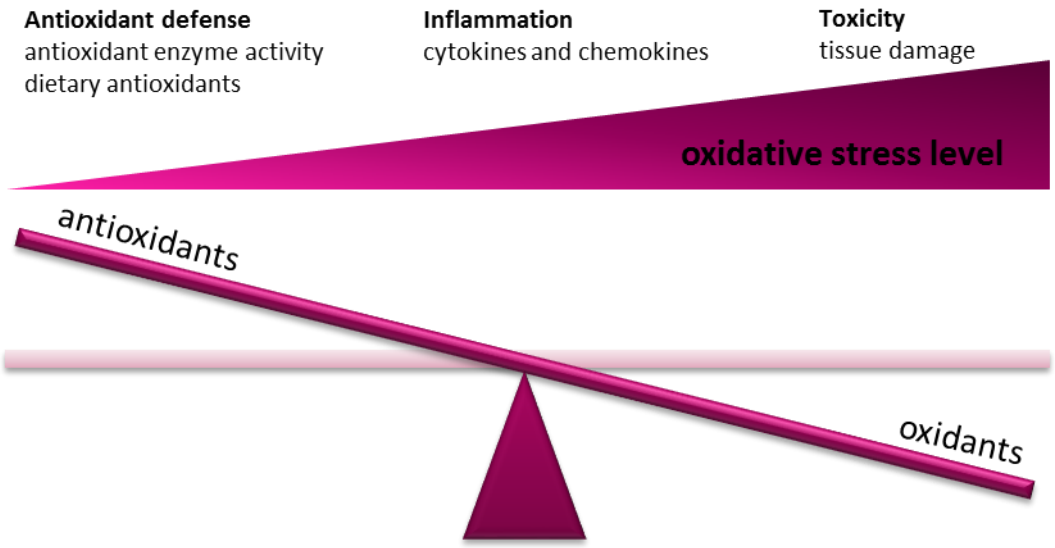

Figure 2.1 Increase in ROS levels leads to an oxidant antioxidant imbalance, initiation of inflammation and oxidative stress-induced tissue damage.

2.3 ENVIRONMENTAL FACTORS

Changes in environmental and lifestyle exposures are probable explanations for the increased prevalence of asthma and allergic diseases. Both indoor and outdoor environmental factors have been reported to influence the risk of disease, as have dietary habits[18, 19], and conditions like overweight and obesity[20].

In 1989, Strachan reported associations between hay fever and family size as well as the number of older siblings in the household[21]. This "hygiene hypothesis" suggested that lack of early childhood exposure to infectious agents increases the risk of allergic diseases, as a Th1 cell response to bacteria and viruses would suppress the Th2 cell response. However, it has later been shown that certain bacterial and viral infections can increase the risk of asthma[22]. It is now believed that a diverse exposure to microbes in early childhood could contribute to protection against both asthma and allergic disease, although timing and exact species remain to be clarified [23]. Growing up on a traditional farm protects against asthma and allergic diseases[24].

The airway epithelium is the first line of response to inhaled environmental exposures. An intact epithelium barrier will reduce allergen passage, protect against infections and enhance the interaction with the innate immune system[23]. Environmental exposures, such as smoking, air pollution or allergens may decrease epithelial barrier function. Several asthma- and

allergy-related genes, such as cytokines, interleukins and toll-like receptors, are expressed in epithelial cells and contribute to barrier function.

Children are considered particularly susceptible to environmental exposures because of their immature immune function, growth and development, and the fact that they eat, drink and breathe more in relation to their body weight than adults do. Genetic predisposition, parental allergic disease, second hand tobacco smoke exposure and prematurity are some of the more established risk factors for childhood asthma and allergic disease[25].

The timing of exposure, during critical periods of organ and tissue development, seems to be of importance for health outcomes later in life. For example, pregnancy and early life exposure to smoking increases the risk of asthma in children[26, 27]. Similarly, pregnancy or early life exposure to traffic air pollution may have lasting impacts on respiratory outcomes later in life[28, 29]. However, since asthma development is a dynamic process throughout childhood, it is also of interest to study exposures later in childhood in relation to disease[8, 30-32].

2.3.1 Indoor mold and dampness

Indoor mold and dampness is a major global problem and an important environmental risk factor for public health[33]. The estimated prevalence of dampness in households ranges from 10-50% worldwide, and at least 20% of households in Europe, Canada and the Unites States have reported signs of dampness[34].

Occurrence depends on climate, air humidity, construction and ventilation of the building as well as the occupant's behavior. Moist indoor conditions allow microorganisms to flourish, resulting in exposure to numerous microbial agents such as spores, cell fragments containing toxins, inflammatory substances and allergens[35]. These agents have been shown to act as irritants, are immunostimulatory and induce inflammatory responses[34].

There are no recommended guideline values for contamination with microorganisms due to a lack of valid quantitative exposure assessment methods. The recommendation today is simply to prevent or remediate indoor mold and dampness[34].

Self-reported visible mold or dampness in the home has shown associations with respiratory symptoms, respiratory infections, asthma exacerbations, new onset asthma and rhinitis in both children and adults[34, 36-39]. In children specifically, visible indoor mold or dampness has been associated with respiratory outcomes such as cough, wheeze, asthma, asthma exacerbations and allergic rhinitis[34, 37, 40]. There are a limited number of studies looking at interactions between mold exposure and oxidative-stress-related genes in association with allergic and respiratory outcomes[41].

2.3.2 Traffic-related air pollution

Outdoor and indoor air pollution constitute together major environmental risk to health, and outdoor air pollution has been estimated to cause 3 million premature deaths annually

worldwide[42]. Outdoor air pollution levels are increasing globally and it has been concluded that 92% of the world population lives in places that exceed the World Health Organization (WHO) air quality guidelines[42].

Air pollution constitutes a complex mixture of compounds with different chemical and biological properties. Their relative concentrations depend on the source and meteorological and geographical conditions.

The components of ambient (outdoor) air pollution are particulate matter (PM), a complex mixture of small particles and liquid droplets, gases such as ozone, nitrogen oxides (NO_x) and carbon monoxide, as well as vapors, such as volatile organic carbons[43]. PM_{10} (particles with an aerodynamic diameter less than 10 µm) can reach the lower airways upon inhalation and include coarse particles (in the size range of 2.5 to 10 µm) and fine particles (less than 2.5 µm in diameter ($PM_{2.5}$)). Ultrafine particles (particles smaller than 0.1 µm in diameter) are proposed to pass the alveolar wall and reach the circulatory system. NO_x comprise nitrogen oxide (NO) and nitrogen dioxide (NO_2). When NO is emitted into the air, it interacts with ozone and oxygen and forms NO_2.

Combustion of fossil fuel is the major process that causes air pollutants. Motor vehicles, aircraft and boat traffic as well as industry, power plants and household heating are contributing sources[43]. For proximity reasons, emissions from road traffic are of major interest to study in relation to human health. NO_x comes primarily from road traffic exhaust. PM_{10} relates to road dust originating from mechanical wear of breaks, clutches and tires, as well as road wear by studded tires, and sanding and salting of roads in wintertime. NO_2 is the most common human-generated form of NO_x in the atmosphere, and the concentrations are strongly correlated with other pollutants. Therefore, it is often used as a surrogate for total air pollution exposure from road traffic in epidemiological research[44].

Oxidative stress and airway inflammation are mechanisms involved in health effects following inhalation of air pollutants[45]. Gases like NO_x have an oxidative effect and particles can contain various substances such as acids, organic chemicals, metals, and soil or dust particles that act as irritants of the airways when inhaled[43, 46]. Lung damage can therefore occur by direct formation of ROS or indirectly via induction of inflammation caused by higher oxidative stress levels.

According to the European Commission and the Swedish air quality standards, the annual mean of ambient air pollution for NO_2 and PM_{10} should not exceed 40 $\mu g/m^3$ [47, 48]. However, there is no evidence for a threshold below which adverse effects do not occur[44]. Epidemiological studies have shown associations between a reduced lung function growth in children and air pollution at levels below these standards indicating that the current guidelines cannot fully protect human health[49]. Further, such guidelines may not apply to particularly susceptible groups.

Different methods can be used to estimate individual exposure to air pollution. Short-term exposures in relation to health effects can be assessed through e.g., chamber experiments or by letting study participants wear measurement equipment. To investigate long-term effects in cohort studies some common methods are to use proxies such as traffic density, proximity to major roads or measured levels of air pollutants at a nearby monitoring station. Methods to estimate further detailed spatial variation in air pollution levels are dispersion modelling relying on measured concentrations at monitoring sites, emission databases and dispersion modelling or land-use regression (LUR) modelling based on monitoring and linear regression modelling of predictor variables.

Traffic-related air pollution exposure is known to have negative effects on children's health. In particular, associations have been reported between air pollution exposure and lung function development in children[49-51]. Traffic air pollution is also known to cause asthma exacerbations and seems to be associated with acute respiratory infections[45]. Studies have shown an association between traffic air pollution and the development of childhood asthma[46, 52-54]. Two meta-analyses reported that early life exposure was of importance for childhood asthma development, and the magnitude of the association increased with increasing age[52, 54]. Another recent systematic review reported positive asthma associations related to black carbon, NO_2, $PM_{2.5}$ and PM_{10} exposure [55]. The potential mechanisms linking air pollution to asthma are, however, not fully understood. Studies of the association between air pollution and sensitization have shown inconclusive results[53, 56].

2.3.3 Dietary antioxidants

In the Nordic Nutrition Recommendations a diet rich in vegetables, fruits, berries, nuts, seeds, and wholegrain is recommended, to promote general health through its content of essential minerals, vitamins and antioxidants[57]. A decreased intake of antioxidants (fruits and vegetables) and n-3 polyunsaturated fatty acids (from oily fish) and an increase in n-6 polyunsaturated fatty acids (margarine and vegetable oils) is hypothesized to increase asthma and allergic disease in children and adults, as such foods have anti-oxidative and anti-inflammatory properties[19, 58]. Dietary antioxidants can decrease oxidative stress by scavenging of free radicals, and are required by the antioxidant defense system in the lung to inhibit damage by ROS[59].

Maternal diet during pregnancy has been suggested to influence the development of asthma and allergic disease in children[60-63]. Duration of breast-feeding is another hypothesized influential factor, although observational studies on breast-feeding have been inconclusive[64]. In addition, there is increasing evidence that the timing of introduction of solid foods (particularly allergenic foods, e.g., egg, peanut and fish) to the diet may influence the development of allergic disease[65]. The child's own diet is also of importance; a low dietary intake of antioxidants has been suggested to influence asthma and allergic disease in children. However, results have been conflicting and few longitudinal studies have been conducted[16, 25, 66, 67]. The child's diet in school age in relation to disease development in adolescence has been less studied.

Antioxidant supplementation has shown beneficial effects on lung function decline in relation to ozone exposure in asthmatic children with genetic susceptibility to oxidative stress[68]. However, clinical trials have not supported the use of individual supplements for prevention or treatment of asthma and allergic disease[16]. On the other hand, consumption of a balanced diet rich in antioxidants has proven favorable in prevention of asthma and allergic disease[67, 69]. Most children in Sweden get enough vitamins and minerals through their food. However, consumption of fruits and vegetables was only 50% of the recommended consumption of 400 grams per day according to a nation-wide nutrition survey performed by the Swedish National Food Agency among 4-, 8- and 11-year-olds[70].

Measuring dietary intake is challenging for a number of reasons. Assessment can be done using open-ended surveys such as dietary recall or records, or by using a food frequency questionnaire (FFQ) with pre-specified response categories. Both methods require that the child and/or the parent can recall what and how much the child has eaten. FFQs are common in epidemiological studies because they are simple, cost-effective and timesaving. Biochemical markers can be used as a surrogate for dietary intake by measurements in, e.g., blood or urine. However, the values obtained cannot be translated into absolute intake due to internal mechanisms such as absorption, metabolism, disease status or homeostatic regulation that can affect the measured level[71].

2.3.3.1 The concept of dietary total antioxidant capacity (TAC)

Total antioxidant capacity (TAC) of diet is considered to be a better estimate than investigating intakes of individual antioxidants from diet. TAC serves to assess the cumulative action of the antioxidant content in a sample[72]. TAC is therefore better at estimating "real life" antioxidant intake, as thousands of compounds are usually ingested simultaneously. Moreover, individual antioxidants, e.g., vitamin C, E, carotenoids, and polyphenols, are highly correlated with one another, therefore, focusing on an individual antioxidant from foods will not ensure that it is the compound of interest that is truly investigated. In addition, there are synergistic effects between compounds which TAC takes into account. Thus, TAC is considered superior because it aims to capture the whole antioxidant network present in foods, taking the synergistic effects into account.

2.3.3.2 Disease-related modification of consumption

Current or previous allergic disease (or parental allergic disease) may cause an individual to change their diet. Children may for example avoid some fruit or vegetables due to allergic symptoms. This will make it more difficult to investigate the true association between exposure and outcome and has to be taken into consideration when studying dietary intake in relation to allergic diseases[73].

2.4 GENETIC FACTORS

2.4.1 Single nucleotide polymorphisms

The genetic code is made up of sequences of the four nucleotides cytosine (C), guanine (G), adenine (A) and thymine (T). The most commonly studied type of genetic variation is differences in individual nucleotides in the deoxyribonucleic acid (DNA), so-called single nucleotide polymorphisms (SNPs). If the SNP sits in a regulatory region, it could affect gene expression, and if it is located in a coding region, it could alter the function of the gene and the gene product (i.e., the protein). Hence, the variant (also called allele) could be associated with disease, which is often investigated by studying the difference in frequency of the allele between cases and controls.

Given the importance of the oxidative stress and anti-inflammatory mechanisms, genes in these pathways have been highly relevant targets for investigation. Susceptibility of the target organ to oxidative injury appears to be dependent on the ability to upregulate protective systems and it is likely that genetic variations could modify the effects of environmental exposures[74].

2.4.1.1 Glutathione S-transferase pi 1 (GSTP1)

Glutathione S-transferase pi 1 (*GSTP1*), located on chromosome 11q13.2, is a candidate gene that has been studied in relation to asthma and allergy due to its link to oxidative stress mechanisms. *GSTP1* codes for a Phase II enzyme, belonging to the GST family of enzymes, which becomes activated at low levels of oxidative stress in order to protect against electrophiles, oxidative stress and ROS, where the GST transferases catalyze glutathione conjunction to detoxify reactive electrophiles[75]. GSTP1 is the predominant cytosolic GST expressed in lung epithelium, which can assist in the reduction of ROS-induced inflammation and tissue injury upon inhalation of toxic substances[76].

GSTP1 genotypes have been associated with asthma and allergy[77]. A meta-analysis of the functional SNP *GSTP1* Isoleucine105Valine (Ile105Val) in relation to asthma revealed a suggestive protective effect of the Val allele, although with large heterogeneity between studies[78]. Further, the *GSTP1* Val allele showed a weak positive association with wheezing, whereas the association was negative for bronchial hyper-responsiveness, and it was concluded that future research should focus on interaction with environmental oxidative exposures. Studies do show that *GSTP1* genotypes could modify the effects of traffic air pollution oxidative environmental exposures in relation to asthma and allergic disease[79-84] as well as lung function[85]. Only one study before our study I had investigated *GSTP1* in relation to mold exposure; it found no interaction between *GSTP1* Ile105Val and mold in relation to early wheeze in children[82].

2.4.1.2 Tumor necrosis factor (TNF)

Tumor necrosis factor (*TNF*), located on chromosome 6p21, encodes a pro-inflammatory cytokine and has been suggested to be of importance in several inflammatory disorders[86, 87].

Several *TNF* SNPs have been investigated in relation to disease, and in particular, the -308G/A polymorphism which has been associated with asthma[86]. Interaction was also found between *TNF* SNPs and early maternal smoking in relation to early childhood wheeze[88]. In addition, interactions between *GST*, *TNF* genotypes and traffic air pollution in association with sensitization have been reported[81].

2.4.1.3 Neuropeptide S Receptor 1 (NPSR1)

Neuropeptide S Receptor 1 (*NPSR1*) (previously named *GPR154* or *GPRA*), located on chromosome 7p14.3, belongs to the G protein-coupled receptor family and is activated by neuropeptide S (NPS)[89]. *NPSR1* was identified as an asthma candidate gene by positional cloning and has been found to be associated with high serum IgE and asthma[90]. Several studies have confirmed these associations[91-98], whereas other studies report no association between *NRSR1* genotypes and asthma or sensitization[99, 100]. Interaction seen between *NPSR1* genotypes and farming lifestyle in relation to childhood allergy warrants of further investigation of this gene in relation to environmental exposure[101].

2.4.1.4 Genome-wide association studies (GWAS)

Additional genes, other than those suggested from linkage studies or related to certain pathways, might be important for disease susceptibility. In the last 15 years, the research field has developed from linkage studies and candidate gene approaches, to more hypothesis-free (or hypothesis-generating) genome-wide studies, where the whole genome is studied in relation to a disease or trait of interest. The human DNA is 99.9% identical between individuals; still, our DNA encompasses several million SNPs. These are located in both coding or non-coding regions and about 3-5% of the SNPs are estimated to be functional, meaning that they may affect the expression or function of genes[102].

In 2001, the Human Genome Project established the first full genome sequence, covering 1.4 million SNPs of the human genome[103]. Later, the international HapMap project genotyped first 1 million SNPs, and, in the second phase, 3 million SNPs in 270 individuals and established linkage disequilibrium (LD) patterns[102]. LD is the non-random association between alleles at different sites on a chromosome. Alleles located close to each other are more likely to be inherited together and create patterns of SNPs called haplotypes. These haplotype blocks were mapped by the HapMap project. By identifying a set of SNPs, researchers can predict the rest of the SNPs in a haplotype using these maps. SNP genotyping arrays, so-called GWAS chips, make use of these specific tag SNPs and thereby the number of SNPs on a chip can be lowered to about 500,000 common variants and still predict up to millions of SNPs by imputation methods. Imputation with a certain reference panel allows meta-analysis of data from different genotyping chips, as the same set of SNPs is added to each study regardless of the genotyping chip used. The 1000 Genomes Project has now sequenced more than 2,000 individuals and characterized 84.7 million SNPs[104]. The establishment of these methods, as well as increased computational power and large consortia, have made GWAS analyses a facilitated tool in performing genetic research.

In the first GWAS on childhood asthma, variants at the ORMDL sphingolipid biosynthesis regulator 3 (*ORMDL3*) and Gasdermin B (*GSDMB*) loci were found to be strongly associated with asthma in children[105], and later in the large-scale GABRIEL consortium, several additional loci were discovered[106]. The association between the 17q21 locus with asthma has been replicated consistently[107]. Variants within the locus have also shown interaction with smoking exposure[108] and rhinovirus bronchitis[109] in relation to early-onset asthma.

Much hope has also been placed in genome-wide interaction studies (GWISs), where GxE is studied in genome-wide data, with the use of more powerful statistical models[110-113]. Individual genes have been studied for GxE[81, 84, 88, 114]. However, few studies have looked at interaction at the genome-wide level and no study has looked at SNPxNO_2 in a genome-wide manner in relation to asthma in children.

To be able to identify disease mechanisms, the function of the gene is of particular interest. It is also important to understand how the gene is regulated, where and when it is expressed, and if there are differences in these patterns between healthy and diseased individuals.

2.4.2 Epigenetics

Epigenetic changes are chemical modifications of DNA. The most studied epigenetic modification is DNA methylation, where a methyl group is attached to a cytosine that sits next to a guanine in the DNA strain, a so-called CpG site. DNA methylation is functionally involved in several regulatory mechanisms such as imprinting, X chromosome inactivation and silencing of repetitive DNA[115]. About 40% of the CpG islands are located near the promoters of human genes and have been frequent targets for DNA methylation analysis due to their functional relevance in gene repression[116]. However, the relationships between DNA methylation and gene regulation have been shown to be more complex.

The genetic sequence is consistent across cells, although different cell-types may have different methylation patterns and therefore it is important to select an appropriate tissue for analysis. Also, most tissue samples consist of heterogeneous cell types - a fact which must be taken into consideration in analysis of DNA methylation levels[117].

Different methods can be used to study DNA methylation: a targeted locus-specific approach, a genome-wide approach or a global methylation approach. The latter provides information about total methylated CpG sites in a sample, whereas no information is given about the status at individual loci[118]. Epigenome-wide association studies (EWASs) can be performed using high-throughput sequencing or microarrays.

DNA methyltransferases (DNMTs) are enzymes that add methyl groups to DNA[115]. Since DNMTs are not active in the DNA amplification step, the methyl groups will then be erased from their positions. Bisulfite treatment is a commonly used method that selectively changes all unmethylated cytosines to uracils (**Figure 2.2**). When sequenced or genotyped a thymine at the position of a CpG site is an indicator of an unmethylated cytosine.

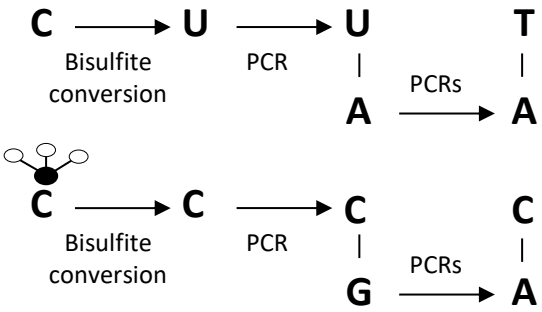

Figure 2.2 After bisulfite treatment, unmethylated cytosine (C) is converted to uracil (U), whereas methylated C is unaffected. During the first round of polymerase chain reaction (PCR) U will pair up with adenine (A) and in later steps A will pair up with thymine (T).

Genetic variability within CpG sites and SNPs located more distantly from the CpG site could affect DNA methylation. Distantly located SNPs that affect DNA methylation are named methylation quantitative trait locus (methQTL). Methylation patterns could also vary based on environmental exposure and have therefore become a hypothesized mediator of the association between environment and gene function.

Environmental factors such as maternal smoking and air pollutants linked to asthma may induce epigenetic effects, although it is not known whether differential DNA methylation is of actual importance in the development of asthma[119, 120]. Lately, childhood asthma and environmental exposure, have been shown to be associated with differential methylation in several gene-specific and genome-wide studies[121-124], and may also be related to ageing[125].

2.4.3 Transcriptomics

During transcription, the genetic sequence is transcribed from DNA into messenger ribonucleic acid (mRNA) molecules. The precursor mRNA contains regions of exons and introns. Introns are removed by RNA splicing and a gene can give rise to several types of mRNA splice variants, depending on the combination of exons in the final transcript. Transcriptomics refers to the study of these transcripts. Much like DNA methylation, transcripts are tissue- and time-specific. The level of expression can be assessed through mRNA sequencing, although microarray-based methods are convenient and relatively cheap in comparison. A tool for identifying genetic variants influencing expression is to perform expression quantitative trait locus (eQTL) analyses investigating the association between a SNP and the expression level of a transcript. Expression levels can also be investigated in association with disease outcomes or exposures of interest.

2.5 CONCEPTS OF GENE-ENVIRONMENT INTERACTIONS

In an epidemiological context, an interaction occurs when two or more factors are dependent on each other, for example when the strength and direction of association between an environmental exposure and an allergic disease differ between the strata of the genetic background in an individual[126]. As discussed above, asthma and allergy are complex diseases

with causal mechanisms involving several single causal components such as genes, the environment and probably timing of exposure.

Studying one genetic and one environmental factor in relation to disease has been commonly used in asthma epidemiology, using a candidate gene approach. In GWISs, statistical power is typically challenging, which calls for international collaborations and alternative statistical approaches. In addition, interactions between several genes and environmental factors may be a necessary causal mechanism for the disease to occur and higher-order interaction between multiple factors is difficult to assess. It is also possible that several different causal mechanisms can lead to the same phenotype. Even if we will not be able to identify all necessary factors for disease initiation by studying simple GxE interactions (i.e. one variant – one exposure interaction), identification of one or more of these factors could possibly lead to new insights into the disease mechanisms[127].

In statistics, interaction refers to the departure from the underlying form of a statistical model[128]. When adding an interaction term to a logistic regression model, the interaction is measured on the multiplicative scale. Presence of interaction means departure from multiplicativity and is a measure of relative risk. An interaction term in a linear regression model means departure from additivity and measures absolute risks.

The statistical power of a study is the probability of rejecting the null hypothesis given that the alternative hypothesis is true, i.e., the capability of observing an effect if the effect actually exists. With higher power, the risk of false negatives (Type II errors) decreases. Power calculations can be used to assess the number of subjects needed in an analysis in order to detect an effect of a certain strength. Power is considered a common problem when performing interaction analyses; even more so when performing genome-wide interaction studies, where adjustment for the high number of tests is necessary. To gain statistical power in the performance of a GWIS, a large sample size and efficient analytical approaches are crucial[112, 113].

Given the complexity of studying interactions between genetic and environmental factors in relation to childhood respiratory and allergic disease and the fact that the underlying mechanism is not completely known, the availability of new techniques, analysis approaches and the construction of large consortia offers new possibilities to study GxE in relation to asthma and allergic disease.

3 AIMS

The overall aim of this thesis was to study the role of genetic polymorphisms and epigenetic alterations in interaction with environmental factors related to childhood respiratory and allergic disease.

Study-specific aims:

I. To investigate if the association between early exposure to mold at home in relation to respiratory and allergic disease in children is modified by the functional *GSTP1* single nucleotide polymorphism Ile105Val.

II. To study DNA methylation in the *NPSR1* gene in relation to asthma and environmental exposures.

III. To identify interactions between genome-wide single nucleotide polymorphisms and air pollution exposure in relation to childhood asthma, with further functional analyses of identified loci at the epigenetic and transcriptomic levels.

IV. To study the association between dietary antioxidant intake and development of asthma and allergic disease, and to assess potential effect modification by known risk factors.

4 MATERIAL AND METHODS

4.1 THE BAMSE BIRTH COHORT

All four studies in this thesis include data from the BAMSE study (Swedish abbreviation for Children, Allergy, Milieu, Stockholm, Epidemiological study). BAMSE is a prospective population-based birth cohort from Stockholm, Sweden[129]. The aim of BAMSE was to study the impact of environmental and life-style factors on asthma and allergy in childhood. Enrollment took place between February 1994 and November 1996. At their first child health visit, parents of all newborn children in four predefined areas of Stockholm County were asked to participate (**Figure 4.1**). The areas represented inner city, urban and suburban areas comprising different types of buildings, traffic air pollution levels and socioeconomic status of parents.

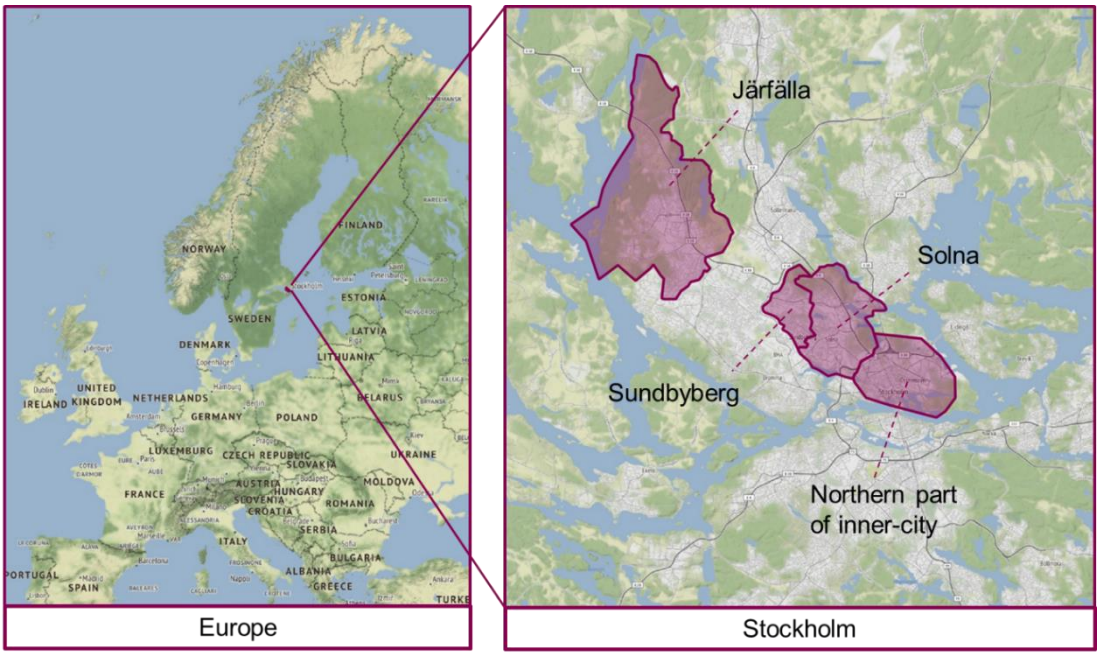

Figure 4.1. The four predefined recruitment areas of Stockholm County, Sweden for the BAMSE birth cohort. Maps adapted from ©OpenStreetMap contributors.

Questionnaires concerning baseline characteristics, such as perinatal conditions, socio-economy and parental allergic diseases, were mailed to the parents when the children were two months old. Out of the 7,221 newborns, 477 had unavailable addresses and 1,256 were actively excluded because the family planned to move within one year of the study start, the parents had inadequate knowledge of the Swedish language, the family had a seriously ill child, or an older sibling already participated in the study. Of the 5,488 eligible children, 502 parents declined participation and 897 never answered the inclusion questionnaire which resulted in a final cohort of 4,089 children. Through the years, questionnaires regarding symptoms of allergic disease, lifestyle factors and major exposures have been filled out by their parents, and the children have participated in clinical examinations including blood

sampling and lung function measurements (**Figure 4.2**). When the children were 4, 8 and 16 years of age, blood samples were analyzed with ImmunoCAP for specific IgE antibodies to common inhalant allergens (cat, dog, horse, birch, timothy, mugwort, *Dermatophagoides pteronyssinus* and *Cladosporium*) using the Phadiatop mix, and to food allergens (cow's milk, hen's egg, codfish, soybean, peanut and wheat) using the fx5 mix[8].

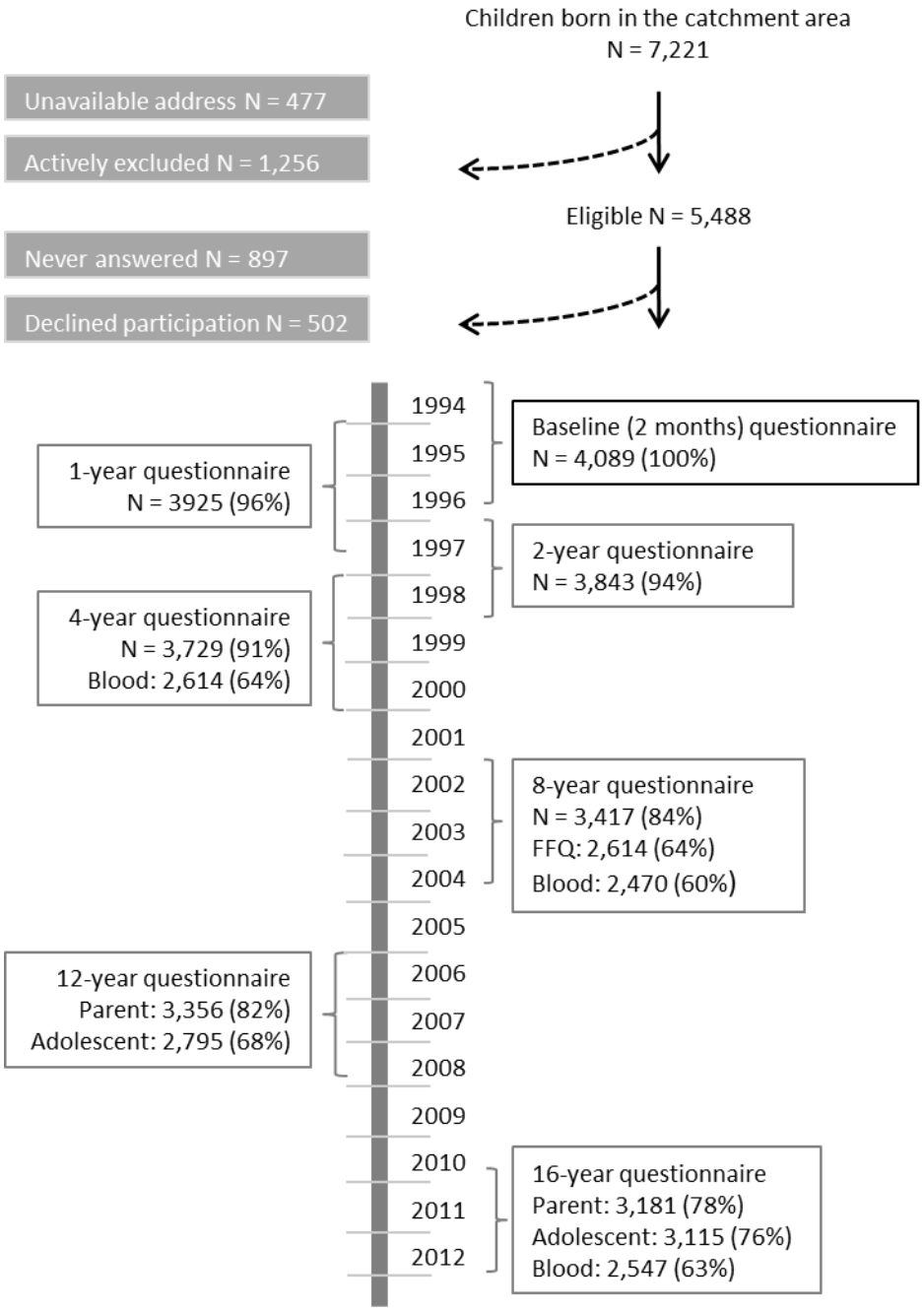

Figure 4.2 Flow chart of the BAMSE birth cohort with initial inclusion and follow-up periods with number of parents (or adolescents) who answered the questionnaire and number of blood samples obtained at the clinical examinations. FFQ: food frequency questionnaire.

DNA was extracted from peripheral blood samples when the children were aged 4, 8 and 16 years. Gene-specific and genome-wide genotyping as well as genome-wide DNA methylation and transcriptome analyses have been performed (see section 4.5.3 to 4.5.6). At the clinical examination at 8 years of age the parents alone or together with the child answered a FFQ about the child's consumption of 98 common food and drink items during the previous 12 months (see section 4.5.7).

Individual estimates of annual air pollution levels have been calculated using both dispersion models and LUR models at home, day care and school addresses from birth up to the 16 year follow-up (see section 4.5.2).

4.2 OTHER COHORTS

In studies **I** and **III** meta-analyses were performed including children from European and North American cohorts as part of the TAG (Traffic pollution, Asthma and Genetics) collaboration. The aim of the TAG collaboration was to assess interactions between traffic air pollution and genetics and the risk of childhood asthma development. The collaboration included, apart from BAMSE, other similar birth cohorts with the aim to study risk factors for asthma and allergic diseases, with available information on individual-level traffic air pollution exposure, allergic disease status and genetics (n = 11,760 children)[130]. The included cohorts were **GINI** (German Infant study on the influence of Nutritional Intervention plus environmental and genetic influences on allergy development)[131], **LISA** (Lifestyle related factors, Immune System and the development of Allergies in East and West Germany plus the influence of traffic emissions and genetics study)[132], **PIAMA** (Prevention and Incidence of Asthma and Mite Allergy)[133], **CAPPS** (Canadian Asthma Primary Prevention Study)[134], **SAGE** (Study of Asthma, Genes, and Environment)[135] and **ALSPAC (**Avon Longitudinal Study of Parents and Children)[136] (**Table 4.1**). In addition, the American childhood asthma study **CHS** (Children's Health Study) was included as a look-up dataset in study **III** (see below).

Study **II** comprises data from the **BIOAIR** (Longitudinal assessment of clinical course and BIOmarkers in severe chronic AIRway disease) cohort. The BIOAIR cohort includes 233 adult patients with severe asthma, mild-to-moderate asthma or COPD, from 12 centers in 9 European countries[137]. The aim was to study severe asthma and the patients were followed up for one year, starting in 2001. Baseline data from BIOAIR, obtained through questionnaires, measurements of lung function and peripheral blood sampling for genotyping and DNA methylation, have been used in study **II**.

CHS, a longitudinal study of several cohorts from Southern California, USA has investigated long-term effects of air pollution on respiratory health in children[138]. In 1993, 1996 and 2003, children aged 5-14 years were recruited from 16 Southern California communities, with follow-up in 2007. Data on GWAS, air pollution and asthma status were used for look-up of main findings in study **III**.

Table 4.1 Overview of the birth cohorts included in this thesis. Acronyms and study areas, study design, recruitment period, the number of subjects enrolled in each cohort at baseline and in studies I-IV

Study	Study design	Recruitment	N baseline	N study I	N study II	N study III	N study IV
ALSPAC, Bristol, UK	Population based birth cohort	1991-1992	14,541	8,639	-	-	-
BAMSE, Stockholm, Sweden	Population based birth cohort	1994-1996	4,089	956	546	481	2,359
CAPPS, Vancouver and Winnipeg, Canada	Birth cohort with asthma intervention for high-risk infants	1995	549	351	-	186 including SAGE	-
GINI, Munich and Wesel, Germany	Population based birth cohort with nutrition intervention	1995-1998	5,991	1,778	-	725 including LISA	-
LISA, Munich, Wesel, Bad Honnef and Leipzig, Germany	Population based birth cohort	1997-1999	3,097	996		see GINI	
PIAMA, The Netherlands	Population based birth cohort with mattress cover intervention and allergic/non-allergic parents	1996-1997	4,146	1,875	-	328	-
SAGE, Manitoba, Canada	Population based birth cohort	1995	16,320	-	-	see CAPPS	-

4.3 FUNCTIONAL STUDIES AND BIOINFORMATICS RESOURCES

Studies **II** and **III** include functional studies with biosamples from adults and data from online resources (in addition to data from BAMSE and other cohorts) as described below.

The **MALF study** in Stockholm provided blood samples from six adult male healthy controls. Previously isolated peripheral blood mononuclear cells (CD4+ T cells, CD8+ T-cells, CD56+ NK cells, CD19+ B cells, CD14+ monocytes) and granulocytes (neutrophils and eosinophils) were investigated for DNA methylation status in the *NPSR1* promoter using the Illumina 450K array (see section 4.5.5) (study **II**)[117].

Nuclear extracts from colon adenocarcinoma cells and human embryonic kidney cells were investigated for nuclear protein binding in relation to genotype and DNA methylation status in *NPSR1* in study **II**, using electrophoretic mobility shift assay (EMSA), in four out of twelve CpG sites in the predicted *NPSR1* promoter.

A data resource of **lung tissue** from 1,111 human subjects who underwent lung surgery at Laval University, University of British Columbia or University of Groningen was used for functional follow-up of GWAS findings in study **III**. The subjects were previously analyzed for whole genome genotyping and genome-wide transcriptomics[139, 140].

Blood samples from sixteen adult non-smokers with asthma and/or airway hyper-responsiveness were previously analyzed for genome-wide DNA methylation pre- and post-**diesel exhaust exposure**[141]. These data were used for functional follow-up of GWAS findings in study **III**.

The **GTEx** (Genotype-Tissue Expression) project is a resource database that enables tissue-specific study of genetic variation in relation to gene expression[142]. Data from 338 blood donors were used for functional follow-up of GWAS findings in study **III.**

FANTOM5 (Functional ANnoTation Of the Mammalian genome) is a database of human cell-type-specific gene expression that was used in study **III** to check gene expression of GWAS findings in different tissues[143, 144].

The **Human Protein Atlas** is a database providing information of protein expression for all major human tissues and organs[145]. It was used in study **III** to identify protein expression in normal respiratory epithelial cells, pneumocytes, lung macrophages and smooth muscle tissue of the lung.

4.4 STUDY POPULATIONS AND STUDY DESIGN

Study **I** included children from the birth cohorts BAMSE, GINI, LISA, PIAMA, CAPPS and ALSPAC with information available on early mold and/or dampness exposure, *GSTP1* (rs1695) genotype as well as respiratory and allergic outcomes up to 8 years of age. In total, 14,495 children had information on genotype.

Study **II** included asthma cases and controls from the 8-year follow-up in BAMSE (n = 546) and adult patients with mild-to-moderate asthma, severe asthma and COPD from BIOAIR (n = 171) with information available on genotype and DNA methylation status in the *NPSR1* promoter. Functional analyses were made in nuclear extracts from colon adenocarcinoma cells and human embryonic kidney cells (EMSA) and in cell-sorted peripheral blood from six healthy male donors (CpG site methylation).

Study **III** included asthma cases and controls from the cohorts BAMSE, GINI, LISA, PIAMA, CAPPS, SAGE and CHS, who were approximately 8 years of age, with GWAS data and individual traffic air pollution exposure (NO_2) data available (n = 3,322). Functional follow-up with DNA methylation and gene expression analysis was performed. Blood samples from 460 8-year-olds in BAMSE and from sixteen adults with asthma and/or airway hyper-responsiveness were investigated for DNA methylation. Blood samples from 263 16-year-olds in BAMSE and lung specimens from 1,111 adults were investigated for gene expression. The online resources GTEx, FANTOM5 and Human Protein Atlas were used (**Table 4.2**).

Table 4.2 Summary of study material and study design for study III

	Material	Association tested	Statistical analysis	Covariates adjusted for
GWIS meta-analysis	BAMSE, GINI/LISA, PIAMA N = 1,534	SNPxNO_2 at birth → asthma[1]	Logistic regression, Two-step approach, 2 df test	Age, sex, city or region at birth, smoking, ancestry
SNPs with interaction p < $1x10^{-4}$ selected for look-up (n = 186 SNPs)				
Look-up	CHS[2] N = 1,602	SNPxNO_2 at 8 y → asthma	Logistic regression	Age, sex, community, smoking, ancestry
	CAPPS and SAGE N = 186	SNPxNO_2 at birth → asthma	Logistic regression	Age, sex, city or region at birth, smoking, ancestry
SNPs with interaction p < 0.05 selected for functional follow-up (n = 8 SNPs)				
Transcriptomics +/-250kb of SNP	Lung tissue N = 1,111	eQTL	Linear regression	Age, sex, smoking status
Transcriptomics +/-250kb of SNP	BAMSE 16 y N = 250	NO_2 at birth → expression at 16 y	Linear regression	Age, sex, peripheral blood cell count[3]
Transcriptomics	GTEx blood donors N = 338	eQTL		
Look-up of GTEx SNPs	BAMSE 16 y N = 173	eQTL	Linear regression	
Gene expression	FAMTOM5			
Protein expression	Human Protein Atlas			
Epigenetics +/-50kb of genes	Adults with asthma and/or airway hyper-responsiveness N = 16	Diesel → methylation	Linear mixed effects modelling	Peripheral blood cell count
Epigenetics +/-50kb of genes	BAMSE 8 y N = 460	NO_2 at birth → methylation	Robust linear regression	Age, sex, smoking, region at birth, asthma[1], blood cell count[4], batch[5]

Smoking: parental-reported environmental tobacco smoke exposure, during the first year of life for all cohorts except CHS and SAGE, which reported at baseline (mean 8.8 years of age) and 8 years of age, respectively. Kb: kilobases. [1]Asthma: doctor's diagnosis of asthma ever, up to 8 years of age. [2]CHS reported all variables at baseline (mean 8.8 years of age). [3]Erythrocytes, platelets, lymphocytes, monocytes, neutrophils, eosinophils and basophils. [4]Estimated counts of CD4+ T cells, CD8+ T cells, NK cells, B cells, monocytes and neutrophils. [5]Bisulfite treatment date.

Study **IV** included all children from BAMSE who had answered the questionnaires at 8 and 16 years, as well as the FFQ at 8 years. Children were excluded if their mean energy intake was below or above 3 log standard deviations (corresponding to ≤ 842.5 or ≥ 3961.1 kcal per day). The study compromised 2,359 children.

4.5 ASSESSMENT OF EXPOSURES AND COVARIATES

4.5.1 Indoor mold and smoking

Information on exposure and covariates that were based on questionnaires answered by the parents at each follow-up, or by adults in the BIOAIR cohort, are listen in **Table 4.3** below.

Table 4.3 Exposure and covariates obtained from questionnaires

Exposures	Definition	Study
Mold	Mold and/or dampness in any room of the home during the first 2 years of life (yes/no)	I
Parental smoking in infancy (birth)	Daily parental tobacco smoke exposure at birth of the child (yes/no)	II, III
Current parental smoking (8y)	Daily parental tobacco smoke exposure at child's age 8 years (yes/no)	II, IV
Smoking status (adults)	Never-smoker, former smoker, current smoker at baseline	II

4.5.2 Traffic-related air pollution

In studies **II** and **IV**, annual averages of traffic-NO_x exposure were estimated using dispersion modelling[146, 147]. In study **III**, annual averages of traffic-NO_2 exposure were estimated using LUR models[148]. For BAMSE, addresses were obtained from the questionnaires and converted into geographical coordinates using a property register managed by the Swedish mapping, cadastral and land registration authority. Missing address information was obtained from the tax authority. Estimated averages were calculated for addresses in Stockholm and Uppsala counties.

4.5.2.1 Dispersion modelling

Information about various sources contributing to measured air pollutants was acquired from an emission database administered by the Stockholm and Uppsala Air Quality Management Association[149]. Road traffic-NO_x was used as a marker for exhaust particles and emission databases for traffic-NO_x used in this thesis were available from 1990, 1995, 2000, 2002, 2003, and 2004[150]. Time- and space-resolved emissions from traffic (including road type, traffic counts, the proportion of heavy traffic and speed limits) at road links were calculated by SLB analysis at the Stockholm Environment and Health Administration using the EVA model[151]. Linear interpolations were made to obtain traffic-NO_x concentrations for all years.

Annual mean concentrations of traffic-NO_x in the study areas were calculated using a Gaussian air quality dispersion model and a wind model, both part of the Airviro Air Quality Management System[152], including meteorological conditions and topography. Concentrations were calculated using a 25 m resolution grid in urban areas. Less populated urban areas and suburban areas used 100 m and 500 m resolution grids, respectively. To compensate for the coarse resolution of the dispersion calculations over suburban areas, additional adjustments were made based on the concentration gradient from roads with more than 10,000 vehicles

per day. The impact of street canyons of the most polluted street segments with multistory houses on both sides in the Stockholm inner city area was accounted for using an Airviro street canyon model[152] for addresses located closer than 30 m to such street segments.

Time-weighted average exposure to NO_x was estimated for each child depending on the duration of time spent at the different addresses. If two home addresses were reported for the same time period, the child was assumed to spend 50% of its time at each place. For the first year of life, only residential addresses were considered since children in Sweden seldom start day care before 1 year of age. Dispersion data was included in studies **II** and **IV**.

4.5.2.2 Land Use Regression modelling

In study **III**, annual average of traffic-NO_2 exposure estimates at birth addresses for all European cohorts were obtained from LUR models developed through the ESCAPE (European Study of Cohorts for Air Pollution Effects) project[148]. In BAMSE, a LUR modelling was also performed for home addresses at the time of the 16-year follow-up.

Measurements of NO_2 were performed between October 2008 and April 2011 at 40 sites within each study area including regional background, urban background and traffic sites. NO_2 was measured for 2 weeks in warm, cold and intermediate seasons and an average annual estimate was calculated. Each study area had a background reference site with continuous measurement of NO_2 that was used to adjust for temporal variation.

LUR models were developed based on the calculated annual concentration and potential predictors from European-wide as well as local Geographical Information System (GIS) databases. Predictors including digital road network, land use, population density, altitude and study-specific local data were added to the model using a supervised forward stepwise procedure, in order to explain the spatial variation of NO_2. One model was developed per study center.

LUR model exposure data for CAPPS and SAGE children born in Winnipeg were based on measurement campaigns from 2007[153]. For CHS, regional air pollution was measured continuously from central site monitors in 16 communities. Annual averages for when the children were 8.8 years were calculated and used for analysis[154, 155].

4.5.2.3 Short-term diesel exhaust exposure

In a double-blind cross-over study, 16 adult asthmatics free from current use of asthma medications were exposed to filtered air or diesel exhaust at a concentration of 300 $\mu g/m^3$ (containing 0.22 ppm NO_2) for 2 hours at two time points, two weeks apart[141]. Throughout the exposure, the participants alternated between light exercise on a stationary bicycle for 15 minutes and rest for 45 minutes.

4.5.3 SNP genotyping

SNPs in **Table 4.4** below were analyzed in studies **I**, **II** and **IV**. In BAMSE, *GSTP1*, *TNF* and *NPSR1* SNP genotyping was performed through matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (Sequenom, San Diego, CA, USA) as earlier described[81, 156]. Genotyping of *GSTP1* in GINI and LISA was performed using the restriction fragment length polymorphism approach[157]. PIAMA and ALSPAC used the Competitive Allele-Specific PCR (KASPar) genotyping chemistry (KBiosciences, Hoddesdon, Herts, UK)[158, 159] and CAPPS used the Illumina BeadArray system (Illumina, San Diego, CA, USA) for *GSTP1* genotyping[96]. In BIOAIR, *NPSR1* SNP genotyping was performed using TaqMan allelic discrimination on the ABI Prism 7500 detection system[156].

Table 4.4 Summary of genes, SNPs and the number of cases and controls from the BAMSE birth cohort that were genotyped for each SNP in studies I, II and IV

Gene	SNP ID	N cases and controls	Study
GSTP1	rs1695 (Ile105Val)	497 with any wheezing up to 4 years of age and 485 never-wheezers	I, IV
	rs762803, rs749174, rs1138272 (Ala114Val), rs1871042, rs4891		IV
TNF	rs1799964, rs1799724, rs1800629 (-308G/A), rs1800610, rs3093664		IV
NPSR1	rs2168891, rs2168890, rs2530547, rs887020	273 with doctor's diagnosis of asthma ever, up to 8 years of age and 273 never asthmatics	II

4.5.4 GWAS

In study **III**, genome-wide SNP genotyping was performed with a chip-based method. The following platforms were used: Illumina 610-Quad BeadChip (BAMSE, PIAMA, CAPPS and SAGE)[106], Affymetrix 5.0 (GINI and LISA)[160], Illumina HumanHap550, HumanHap550-Duo or Human610-Quad BeadChip (CHS)[161], Illumina Human1M-Duo BeadChip (Lung tissue data) [139]. Imputation based on HapMap 22, NCBI build 36, resulted in more than 2 million overlapping SNPs between the discovery cohorts. Details on quality control (QC) of samples are provided in the original manuscripts.

4.5.5 Epigenetics

For study **II**, peripheral blood cell DNA was bisulfite-treated using EZ-96 DNA Methylation Kit (Zymo Research Corporation, Irvine, CA, USA) and analyzed for DNA methylation in the predicted *NPSR1* promoter using EpiTYPER[162]. Five CpG sites from 8-year-olds in BAMSE and 12 CpG sites from adults in BIOAIR were assessed. For the six healthy male donors, DNA was bisulfite-treated and analyzed on the Illumina Infinium HumanMethylation450 BeadChip (Illumina 450K) (Illumina Inc., San Diego, CA, USA) as well as using EpiTYPER[117].

Blood samples from 460 8-year-olds with asthma and controls in BAMSE and from sixteen adult non-smokers with asthma and/or airway hyper-responsiveness were assessed for whole genome methylation using bisulfite treatment and Illumina 450K[123]. In the sixteen adults, blood was collected pre- and 6 h as well as 30 h post-exposure to filtered air and diesel exhaust[141]. Details on QC of samples are provided in the respective original manuscripts and the original manuscript for study **III**.

4.5.6 Transcriptomics

In study **III**, blood samples from 263 16-year-old participants in BAMSE were investigated for gene expression using the Affymetrix Human Transcriptome Array 2.0 Genechips (Affymetrix, Inc., Santa Clara, CA, USA) (173 of these had GWAS data available and 250 had NO_2 data available)[123]. The lung tissue data from adults was obtained through whole-genome expression profiling using an Affymetrix custom array (see GEO platform GPL10379, http://www.ncbi.nlm.nih.gov/geo/)[139, 140]. Details on QC of samples are provided in the original manuscripts.

4.5.7 Food frequency questionnaire

A FFQ was filled out by the parents alone or together with the child at the 8-year clinical examination within BAMSE (n = 2,614). The FFQ contained questions about the average consumption during the previous 12 months of 98 foods and beverages frequently consumed in Sweden. There were ten pre-specified response categories ranging from 0 times/month to ≥ 3 times/day. The FFQ has been transformed into individual energy and nutrient intake using composition values obtained from the Swedish National Food Administration Database[163]. Variability in nutrient intake caused by differences in total energy consumption was managed by energy-adjusting the calculated nutrient levels using the residual method[164].

The concept of TAC was used to assess children's total dietary antioxidant intake (study **IV**). Estimated dietary TAC values (μmol Trolox equivalents/g (TE/g)) for food and drink items included in our FFQ were derived from a United States (US) database[165]. One hundred foods and drinks commonly consumed in the US had been assessed for their TAC content using the oxygen radical absorbance capacity (ORAC) assay[166]. Of these 100 items, 35 were available in our FFQ (**Table 4.5**). The amount of intake of each food item in the FFQ was multiplied with its respective estimated TAC value. Total dietary TAC (μmol TE/day) was obtained by summarizing TAC values for all food items eaten and energy-adjusting using the residual method[164].

Table 4.5 Food items from the FFQ at 8 years with available ORAC-TAC estimates that contributed with > 1% to the total ORAC-TAC in study IV (19 of 35 food items)

Food item	Average proportion of ORAC-TAC (%)[1]
Apple, pear	23.3%
Juice	10.3%
Chocolate milk	9.1%
Boiled potato	7.5%
Other fruit (apricot, kiwi, nectarine, peach, plum, grapes)	6.1%
Orange, citrus fruit	4.5%
Stewed fruit, fruit soup	4.0%
Rye bread	3.3%
Berries (blackberry, blueberry, raspberry, strawberry)	3.2%
Banana	2.8%
Mashed potato	2.7%
Whole meal bread	2.6%
Tea	2.5%
Carrot (raw, canned, cooked)	2.2%
Breakfast cereal	2.2%
Oatmeal	2.1%
Chocolate	2.0%
Jam, marmalade	1.5%
Crisp bread	1.4%

[1]Calculated by combining the estimated ORAC-TAC value and the frequency of consumption.

4.5.8 Other characteristics

Sampling season: the time of year when blood samples were collected at the 8-year clinical examination was used in study **II** and divided into three-month intervals for analysis: January-March (reference group), April-June, July-September, and October-December.

Body mass index: body mass index at 8 years of age in BAMSE and at baseline in BIOAIR was measured in kg/m^2 based on height and weight information (study **II**).

4.6 DEFINITION OF HEALTH OUTCOMES

4.6.1 Asthma and allergy-related health outcomes

The asthma and allergy-related health outcomes were assessed through questionnaires filled out by parents at each follow-up (**Table 4.6**).

Table 4.6 Definitions of health outcomes

Variable	Definition	Study
Children		
Early wheezing (0-2 y)	Any parental-reported wheezing or whistling in the chest up to two years of age.	I
Early asthma (0-3y) School-age asthma (6-8 y) Asthma (16 y)	Parental-report of at least two of the following three conditions: ever doctor's diagnosis of asthma, wheezing in the last 12 months and asthma medication in the last 12 months. Teenagers' own reports were used at 16 y.	I (0-3 & 6-8 y) IV (8 & 16 y)[1]
Current asthma (8 y)	Parental-reported ever doctor's diagnosis of asthma up to 8 years in combination with at least one episode of wheeze in the last 12 months.	II
Allergic asthma (8 y)	Parental-reported ever doctor's diagnosis of asthma up to 8 years in combination with an IgE-value for inhalant allergens ≥ 0.35 kU/l at 8 y. Controls were children with no history of asthma or other allergic diseases.	II
Non-allergic asthma (8 y)	Parental-reported ever doctor's diagnosis of asthma up to 8 years in combination with an IgE-value for inhalant allergens < 0.35 kU/l at 8 y. Controls were children with no history of asthma or other allergic diseases.	II
Ever doctor's diagnosis of asthma (0-8 y)	Parental-reported ever doctors doctor's diagnosis of asthma (up to approximately 8 years of age). CAPPS: clinical examinations by a pediatric allergist at 7 y. SAGE: parental-reported doctor's diagnosis of asthma at 7 y with confirmation of the diagnose by a pediatric allergist (mean age 9y). CHS: parental-reported (or self-reported) doctor's diagnosis of asthma at baseline (1994-2003, age range 5-14 years) or at any time during the follow-up (until 2007).	II III
Allergic asthma (16 y)	A combination of asthma (16 y) and sensitization (16 y) against any inhalant allergen.	IV[1]
School-age nasal symptoms (6-8 y)	Parental-reported prolonged sneezing, runny or blocked nose without common cold or flu in the last 12 months. CAPPS: Physician-diagnosed rhinitis.	I
Ever nasal symptoms (2-10 y)	Same as School age nasal symptoms 6-8 y.	I
School-age rhino-conjunctivitis (6-8 y)	Parental-reported runny, itchy eyes in addition to nasal symptoms.	I
Rhinitis (8 or 16 y)	Parental-reported prolonged sneezing, runny or blocked nose without common cold or flu in the last 12 months, and/or nose or eye symptoms after contact with furred pets and/or pollen after 4 years of age for rhinitis at 8 years of age, or after contact with furred pets and/or pollen in the last 12 months for rhinitis at 16 years of age. Teenagers' own reports were used at 16 y.	IV[1]
Allergic rhinitis (16 y)	A combination of rhinitis (16 y) and sensitization against any inhalant allergen (16 y).	IV[1]

Adults		
Mild-to-moderate asthmatics	Patients had stable disease and received daily treatment with a maximum of 800 µg/day budesonide or beclomethasone, 500 µg/day fluticasone or equivalent. The patients used short-acting β-agonists as needed but did not require treatment with long-acting β-agonists and had had neither exacerbations nor hospitalizations in the past year.	II
Severe asthma	Specialist treatment for at least 1 year and at least one exacerbation requiring oral steroid treatment in the past year, despite continuous treatment with high doses of inhaled corticosteroids (ICS; at least 1,600 µg/day budesonide or beclomethasone, 800 µg/day fluticasone or equivalent), and long-acting β-agonists or oral theophylline for at least 1 year.	II
COPD	Physician-diagnosed chronic obstructive pulmonary disease.	II
FEV_1	Forced Expiratory Volume in 1 second (the air exhaled in one second in spirometry test at baseline).	II
FVC	Forced Vital Capacity (the total volume of air exhaled in spirometry test at baseline).	II

[1]Onset of allergic disease between 8 and 16 years of age was defined as fulfilling the definition at 16 years but not at 8 years of age.

4.6.2 Sensitization

Sensitization to inhalant allergens in BAMSE was defined as a specific IgE result of ≥ 0.35 kU_A/L against cat, dog, horse, birch, timothy, mugwort, *Dermatophagoides pteronyssinus* or *Cladosporium* (at 8 years in study **I**, **II** and **IV** and also at 16 years in study **IV**). In study **I**, sensitization to inhalant allergens at 6-8 years of age in BAMSE, GINI, LISA and PIAMA was defined as a specific IgE result of ≥ 0.35 kU_A/L against at least one of the measured inhalant allergens in each separate cohort. Sensitization in ALSPAC and CAPPS was assessed by skin prick test (SPT) with a positive reaction of having a wheal diameter of > 2 mm to at least one of the measured inhalant allergens. Allergic disease was defined as a combination of the health outcome and sensitization according to the definitions above.

4.7 STATISTICAL ANALYSES

For study **I**, main effects of early exposure to mold at home as well as *GSTP1* polymorphism rs1695 in relation to respiratory and allergic disease in children were assessed using **logistic regression** analysis, adjusting for sex, parental allergy and maternal smoking during pregnancy (although CAPPS did not adjust for smoking due to lack of information). Interaction between early exposure to mold and the *GSTP1* polymorphism was tested by inclusion of an **interaction term** in the model and **stratification** by genotype. Meta-analysis with random effect model was used to account for heterogeneity between cohorts.

In study **II**, asthma status and environmental exposures in relation to DNA methylation in *NPSR1* was assessed using **linear regression** analysis, adjusting for age and sex as well as country of origin in adults and batch effect in children. Results were presented as beta-values per unit change in the exposure level, with corresponding p value. Traffic-NO_x estimates in BAMSE were assessed by the 5th to 95th percentile difference in NO_x exposure levels, corresponding to 45.7 μg/m^3 during the first year of life, and 26.0 μg/m^3 for current exposure at eight years of age. Analyses of air pollution were additionally adjusted for municipality at birth. If a CpG site coincided with a SNP causing or destroying a CpG site the analysis was restricted to homozygous carriers of the allele creating a CpG site.

Statistical analyses for study **III** are summarized in **Table 4.2**. Interactions between genome-wide SNPs and NO_2 exposure (per 10 μg/m^3 increase) in relation to childhood asthma were assessed using **logistic regression**. Fixed effect model with sample size weighted analysis was used in the GWIS meta-analysis. **Robust linear regression** was used to account for influential outliers and potential heteroscedasticity in the methylation data in relation to NO_2 exposure data. Since methylation data were obtained from whole blood, adjustment for cell type composition was performed using the Houseman reference-based method[167]. **Linear regression** was applied for the lung eQTL analyses. Meta-analysis using a fixed effect model was performed with inverse variance weighting. The **linear mixed effects model** is a statistical model used for repeated measures that contain fixed effects and random effects. It was applied to the measurements comparing post-diesel exhaust particles (DEP) vs. pre-DEP, and post-filtered air particles (FAP) vs. pre-FAP in order to identify CpG sites that were differentially methylated as a result of diesel exhaust exposure but not filtered air. **Log-linear regression** on NO_2 exposure in relation to gene expression values was performed in BAMSE and adjusted for cell counts. **Log-linear regression** was also used to analyze eQTLs in BAMSE. Correction for multiple testing was done to limit false positive occurrences. For Bonferroni correction the significance level of 0.05 was divided by the number of hypotheses tested in order to obtain an adjusted significance threshold. The false discovery rate (FDR) significance threshold of 5% was calculated by ranking the p values from lowest to highest and multiplying the rank by 0.05, and the product was divided by the number of tests[168].

In study **IV**, **logistic regression** analysis was performed to investigate the association between dietary antioxidant intake (analyzing TAC in tertiles) and development of asthma and allergic disease, and to assess potential effect modification by known risk factors. Traffic-NO_x at 8 years was analyzed in tertiles, corresponding to median (min-max) values of 3.09 (0.40-4.81), 8.24 (4.82-12.16) and 18.49 (12.21-56.04) μg/m^3 for tertiles 1 to 3, respectively. Sex, parental history of allergic disease, early maternal smoking and parental socioeconomic status (blue-collar or white-collar worker) were adjusted for. **Test for trend** was performed by assigning the median value of dietary TAC within each tertile and tested as a continuous variable in the model.

4.7.1 Tests for interactions

Logistic regression with inclusion of an **interaction term** in the model was used in studies **I** and **III** and a p value for interaction was obtained. **LRT** between the model with and without interaction term was used to test the null hypothesis of no interaction in study **IV**. **Stratified analyses** were performed in studies **I**, **III** and **IV** to investigate the direction of effect over strata.

The **two-degree-of-freedom (2 df) test**[112] (study **III**) jointly tested the main effect of SNP and the SNPxNO_2 interaction effect by logistic regression, in the context of GWAS. Using this test, SNPs with main effects and SNPs with heterogeneity in genetic effects over different levels of NO_2 exposure will be detected. The p values from the separate k ($k = 3$) cohorts were meta-analyzed using Fisher's method, which provided a summarized χ^2 statistic with 2k degrees of freedom[169].

The first step in the **two-step approach**[113] (study **III**) tested the association between SNPs and NO_2 in cases and controls combined by linear regression. All SNPs with a p value < 0.05 were in a second step tested for SNPxNO_2 interaction using logistic regression. A summary of the two-step approach is presented in **Table 4.7**. Step one is based on the case-only method, where associations between genetic variants and environmental exposure are tested in cases only, using linear regression. The method relies on the assumption that the genetic variants and the environmental exposure are independent in the source population and an association between the genetic variant and the environmental exposure will arise among cases whenever GxE is present. Analyzing cases only in step one will produce correlations between step one and step two test statistics, resulting in an inflated number of false positives for the whole test. Therefore, in the current two-step approach, cases and controls were analyzed together in step one. The two-step approach is a more powerful method than testing for GxE in a case-control setting.

Table 4.7 The two-step approach

	Statistical model	Pros	Cons
Step 1	Linear regression: SNP - NO_2 association in cases and controls combined	Uses a powerful "case-only" method that increases the chance of identifying true positives.	Sensitive to the independence assumption with risk of false positives.
		Screening step to reduce the multiple testing burden in step 2.	
Selection of SNPs with p < 0.05			
Step 2	Logistic regression: SNPxNO_2 in relation to asthma using a standard case-control setting	Insensitive to the independence assumption. The bias of false positives in step one is overcome.	

4.8 ETHICAL APPROVALS

The BAMSE study was approved by the Ethics Committee of Karolinska Institutet, Stockholm, Sweden. The parents of the children received written information about the purpose of the study and all study participants have given their written informed consent. The additional respective studies included in studies I-III received ethical approval by their local ethical review boards and subjects provided their informed consent.

5 RESULTS AND DISCUSSION

5.1 GENE-ENVIRONMENT INTERACTIONS IN RELATION TO ASTHMA AND ALLERGY

In study **I**, a meta-analysis of 14,595 children, the association between early exposure to mold at home and respiratory and allergic disease was investigated. The main effect analyses of early mold exposure showed significant associations with early wheezing and ever nasal symptoms during childhood, OR: 1.27 (95% CI: 1.09-1.47) and 1.19 (95% CI: 1.02-1.38), respectively, although with indicated heterogeneity between the cohorts (p value for heterogeneity < 0.05). We further evaluated potential interaction between early exposure to mold and the *GSTP1* polymorphism rs1695. We did not find any statistically significant interaction with respect to allergic disease outcomes. However, stratification by genotype revealed statistically significant associations between mold exposure and school age nasal symptoms as well as ever nasal symptoms in the strata of Ile/Ile carriers (OR: 1.20 (95% CI: 1.00-1.44) and OR: 1.30 (95% CI: 1.04-1.61), respectively), but not in the other genotype group (**Figure 5.1**).

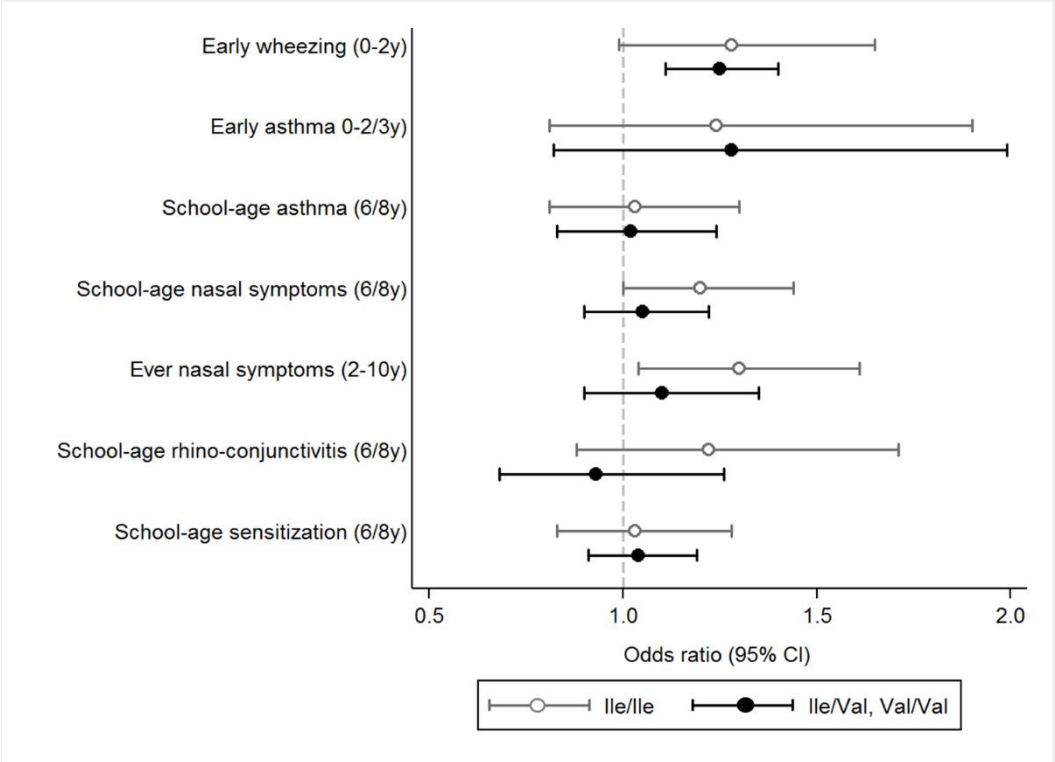

Figure 5.1 Adjusted odds ratios and 95% confidence intervals for the association between early mold exposure and wheezing, asthma, nasal symptoms and sensitization to inhalant allergens, stratified by genotype using dominant coding.

Systematic reviews and meta-analyses[34, 36-40], as well as prospective birth cohort studies[170-173], have found associations between early life exposure to mold and dampness, and wheezing, asthma and rhinitis in children. In line with our results, exposure to visible mold or dampness

has not been associated with sensitization, as evaluated in several prospective birth cohort studies[34, 171-173].

Mold exposure during first year of life has been significantly associated with increased risk of wheezing at 2 years of age, and similar to our study no difference over strata of *GSTP1* genotype was seen[82]. A cross-sectional Taiwanese study did not find interactions between mold exposure and *GSTP1* in 12-year-olds with asthma and controls[41]. Val105 genotype in combination with early mold exposure has been associated with current atopic dermatitis in 3-18-year-olds[174]. In a previous meta-analysis on 31,742 children[173], mold exposure was associated with nasal symptoms, and these associations were also seen in Ile/Ile carriers in the present study. Our study is the first and only collaborative study that has looked at GxE in relation to mold and respiratory and allergic health outcomes using birth cohort data.

Other genes, such as *CHIT1*, *TNF*, *IL13* and *IL4*, which were not tested in the current study, have also shown interaction with mold and dampness in relation to wheeze, asthma and allergic rhinitis[175-178]. In summary, in line with other studies, early mold exposure was associated with nasal symptoms and early wheezing in children. A positive significant association of nasal symptoms in relation to mold exposure was seen primarily in Ile/Ile carriers of *GSTP1* Ile105Val polymorphisms. However, there seems to be a complex interplay between *GSTP1* genotype and environmental exposures and interactions could not be ruled out.

In study **III**, we aimed to identify genetic regions of importance in the association between traffic air pollution and childhood asthma. GWISs of SNPs and traffic NO_2 in relation to childhood asthma up to 8 years of age were performed in three European birth cohorts (n = 1,534). After meta-analysis, several suggestive loci were identified, although no SNP reached a genome-wide significance level, $p < 7.2 \times 10^{-8}$. Two alternative statistical approaches, the two-step approach and the 2 df test, with potentially better power for genome-wide interaction analyses, were also used (see below).

Since the primary GWIS approach was used as a screening tool to identify SNPs and genes that may interact with air pollution in asthma, all SNPs with an interaction p value $< 1 \times 10^{-4}$ in the primary GWIS approach (n = 186 SNPs) were considered for look-up in CHS (n = 1,602) and CAPPS/SAGE (n = 186) (**Figure 5.2**).

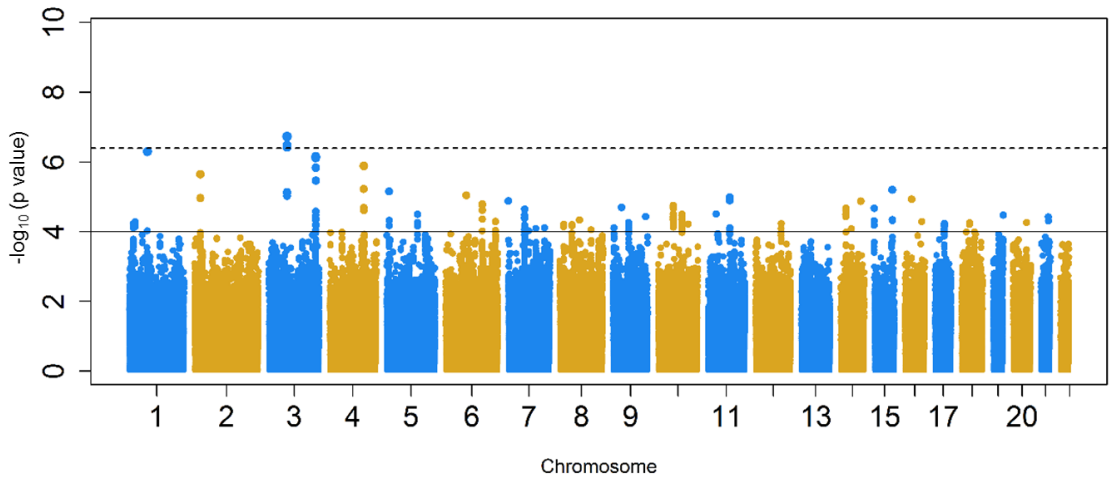

Figure 5.2 Manhattan plot for the association between SNPxNO_2 and asthma from the discovery genome-wide interaction meta-analysis. The lower horizontal line indicates the threshold for bringing SNPs forward to look-up ($p < 1x10^{-4}$). The upper horizontal dashed line indicates the genome-wide significant threshold using the 2-step approach ($p < 4x10^{-7}$). NO_2 is per 10µg/m^3 increase.

Eight SNPs in the loci of adenylate cyclase 2 (*ADCY2*), UDPGal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5 (*B4GALT5*), discs large homolog 2 (*DLG2*) and molybdenum cofactor sulfurase (*MOCOS*) had interaction p values < 0.05 in the larger CHS cohort. Three SNPs located within *ADCY2* on chromosome 5 (LD $r^2 = 0.93$-1.0) showed same direction of interaction effect and same direction in stratified analyses of the association between NO_2 exposure and childhood asthma, when the discovery meta-analysis was compared with CHS (data shown for rs4143882 in **Table 5.1**).

Table 5.1 Adjusted odds ratios and 95% confidence intervals for the association between NO_2 and asthma, stratified by rs4143882 genotype (*ADCY2*) using dominant coding.

	Discovery meta-analysis n = 1,534		**CHS n = 1,602**	
	OR (95% CI)	**p**	**OR (95% CI)**	**p**
GG	0.81 (0.33-1.99)	$4.75x10^{-5}$	0.88 (0.76-1.02)	0.015
AG/AA	1.61 (1.04-2.51)		1.13 (0.98-1.29)	

p: interaction p value using additive coding of SNP. NO_2 is per 10 µg/m^3 increase.

SNPs in *ADCY2* have been associated with pulmonary function in adults[179] and a previous study in BAMSE's 8-year-olds reported an association with lung function for these SNPs with a suggested interaction with current tobacco smoke exposure[180]. One SNP in *ADCY2*[181] and one in *DLG2*[182] has also been associated with COPD. Persistence of childhood asthma into adult life is associated with lung function impairment and increased risk of COPD, a disease which may have its origin early in life[183].

The two-step approach revealed four genome-wide significant SNPs, rs7651862, rs11706125, rs11718057 and rs13066946 close to the membrane-associated guanylate

kinase, WW and PDZ domain containing 1 (*MAGI1*) gene (**Figure 5.2**). These SNPs had the lowest interaction p value in the primary GWIS meta-analysis and they were also nominally significant in the 2 df test ($p < 0.05$). However, the SNP x air pollution interaction was not statistically significant in a look-up in the cohorts CHS and CAPPS/SAGE.

To date, no other interaction study between GWAS and traffic air pollution in association with childhood asthma has been published. Many previous GxE attempts using GWAS data have failed to reach genome-wide gene-by-environment significance for respiratory diseases and lung function. The first GWIS on childhood asthma and tobacco smoke exposure did not reveal any genome-wide statistically significant SNPs[184]. However, suggestive evidence for interaction was found for *EPB4IL3* (interaction $p = 4.29 \times 10^{-5}$ in discovery phase and replication interaction $p = 0.03$ for interaction with *in utero* tobacco smoke exposure) and *PACRG* (interaction $p = 1.37 \times 10^{-5}$ and interaction $p = 0.06$ in discovery and replication phase respectively for interaction with childhood tobacco smoke exposure). No genome-wide significant interaction could be identified for common SNP variants (interacting allele frequencies from 30 to 70%, for which their study was well powered) and farming lifestyle in a GWAS of 1,708 children investigating asthma and atopy[185]. Rare variants, particularly of SNPs related to the *GRM1*, belonging to the G protein coupled receptor family, showed interaction with farming environment (Bonferroni corrected interaction $p < 0.05$). GWISs on smoking and lung function, studying > 50.000 subjects, have not reached genome-wide significance for interaction, although three novel regions of potential importance for lung function were identified using the joint 2 df test in the CHARGE and SpiroMeta studies[186, 187]. In a study investigating interaction with sex in relation to childhood asthma, functional follow-up analyses revealed expression in alveoli and macrophages of the suggested gene *DMRT1* (interaction $p = 5.21 \times 10^{-6}$), a gene required for testicular development[188]. Similarly, a GWAS on bronchial hyper-responsiveness in asthmatics identified no genome-wide significant hits, although subsequent lung eQTL analyses of the top hit SNPs revealed novel genes associated with asthma[189].

There are a few examples of genome-wide significant interaction effects in the field of lung function and asthma, one from a Dutch study of 12,400 adults, where interactions between seven SNPs and occupational exposures to dust gases and fumes in relation to lung function were reported[190]. Two of the SNPs were identified as lung eQTLs for *TMEM176A* and *PDE4D*. In addition, a GWIS on lung function decline in adults identified *CDH13*, a gene functionally linked to the adipokine adiponectin, to interact with 11-year cumulative PM_{10} exposure (interaction $p = 8.8 \times 10^{-10}$), using a two-stage GWIS analysis approach[191]. In a recent GWIS in Puerto Rican children with asthma, significant interaction was detected for dust mite allergen exposure at home and a SNP close to a transcription factor binding site involved in IL-17 signaling pathway in relation to lung function (interaction $p = 3.1 \times 10^{-8}$)[192].

In conclusion, in study **III** we found that interaction between SNPs and air pollution exposure is important for asthma development. Based on our result and previous GWIS attempts by other researchers, we conclude that using a genome-wide significance threshold may be a too

conservative cut-off for look-up of GWIS top hit findings, resulting in false negative results and missing of important loci involved in asthma development.

In study **IV**, the association between dietary antioxidant intake (assessed as dietary TAC) and development of asthma and allergic disease was assessed in the BAMSE birth cohort (n = 2,359). The main effect analyses showed significant inverse association for new onset sensitization to inhalant allergens between 8 and 16 years of age (OR: 0.73 (95% CI: 0.55-0.97) for the 3rd tertile as compared with the 1st, p for trend = 0.031) and for new onset of allergic asthma between 8 and 16 years of age (OR: 0.57 (95% CI: 0.34-0.94) for the 3rd tertile compared with the 1st, p for trend = 0.029). After additional adjustment for dietary intake of specific nutrients, supplement use, and allergic disease in early childhood, the results remained significant for sensitization to inhalant allergens. The significant association with sensitization to inhalant allergens remained also after consideration of disease-related modification of consumption.

In systematic reviews and meta-analyses, a Mediterranean diet, as well as diets rich in fruits, vegetables and vitamins have shown inverse associations with allergic disease outcomes in children[66, 67, 193]. However, a majority of previous studies have been cross-sectional and inconsistent results between studies have been seen.

Inverse associations were detected between fruit intake and current rhinitis in children[73], as well as between beta-carotene and rhinitis[194] in two previous cross-sectional studies. However, after adjustment for disease-related modification of consumption, all associations became weaker and significance disappeared. These studies indicate the importance of taking the concept of disease-related modification of consumption into consideration when studying associations between diet and allergic disease outcomes.

Recent cross-sectional studies in children found inverse associations between dietary intake of vitamin C and E and asthma[195], and inverse associations between intake of vitamin C and allergic rhinitis[196]. No associations were seen for retinol, beta-carotene, vitamin C or vitamin E in relation to sensitization[196]. An antioxidant-rich diet was inversely associated with childhood asthma, but not with rhinitis or sensitization in another cross-sectional study[197]. A longitudinal study found an inverse association between dietary beta-carotene and allergic sensitization in children[198]. However, none of these studies took disease-related modification of consumption into consideration. A longitudinal study of serum concentrations of carotenoid and tocopherol found no evidence of decreased intake of antioxidants in relation to increased asthma prevalence in children[199].

In study **IV**, significant interaction between dietary TAC and current traffic air pollution exposure (NO_x) was observed in relation to sensitization to inhalant allergens between 8 and 16 years of age (interaction p = 0.03), with stronger association among children with low levels of traffic air pollution exposure (**Figure 5.3**). The association was also statistically significant in children who had not been exposed to parental smoking, although no interaction was detected (interaction p = 0.85).

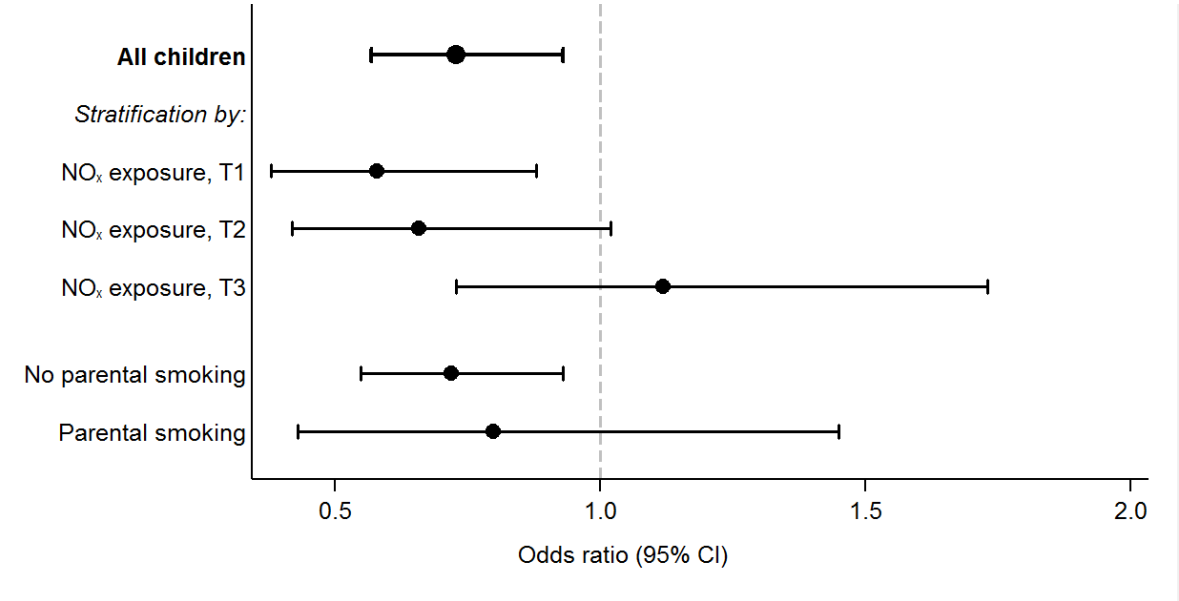

Figure 5.3 Adjusted odds ratios and 95% confidence intervals for the association between dietary TAC intake (first tertile vs. second and third tertiles combined) and sensitization to inhalant allergens. Main effect analysis for all children and stratified analyses by traffic-NO_x exposure in tertiles as well as parental smoking.

One previous cross-sectional study in children found inverse associations for dietary TAC in relation to current asthma, and, much like in our study, the effect was only seen in children with non-smoking parents[200]. In adults, dietary TAC had a favorable role in lung function, in premenopausal never-smoking women[201]. A study on lung function in children found interaction between ozone and high intake of fruits and vegetables and adherence to a Mediterranean diet, suggesting that a high intake could modulate adverse effects by air pollutants[202].

Tobacco smoke exposure in children with a low dietary vitamin A intake was associated with asthma diagnosis in a cross-sectional study[203]. Further, homozygous carriers of the *GSTP1* Ile105 allele showed an even larger effect estimate by exposure. Lower lung function in children with asthma was reported in carriers of a combined effect of *GSTM1*, *GSTP1* and *NQO1* polymorphisms, in combination with a low dietary vitamin C intake[204]. In our study **IV**, we found no significant interactions between dietary TAC and *GSTP1*- or *TNF*-genotypes, although the number of children with both TAC estimates and genotype information available was low and interactions cannot be ruled out. Our results should therefore be interpreted with caution. In summary, the finding from the stratified analyses in study **IV** indicates that dietary antioxidant intake may protect against allergic sensitization in conditions of low oxidative stress.

5.2 EPIGENETIC DIFFERENCES IN ASSOCIATION WITH ASTHMA AND ENVIRONMENTAL EXPOSURES

In study **II**, DNA methylation in the *NPSR1* promoter was investigated in relation to asthma and environmental exposures in the adult BIOAIR cohort (n = 171), as well as in the BAMSE birth cohort (n = 546). Adult severe asthma was inversely associated with DNA methylation at CpG site 5 in the promotor region of *NPSR1* as compared with in mild asthma (-3.29%, p = 0.0001). In children, allergic asthmatics compared with never-asthmatics had a 0.6% lower level of DNA methylation at both CpG site 1 (p = 0.01) and CpG site 4 (p = 0.03), whereas no such associations were found for non-allergic asthmatics. Others have not studied differential DNA methylation in *NPSR1,* but *NPSR1* variants have been associated with sensitization and allergic asthma in several studies[90-92, 98, 205]. Gene expression of *Npsr1* was upregulated in the lung of sensitized mice compared with non-sensitized mice after experimentally induced lung inflammation[90]. *NPSR1* polymorphisms have also been associated with allergic rhinitis[206] and gene-gene interaction between *NPSR1* and *FGF1* genotypes was found in one study in relation to sensitization[207]. EWASs have detected differential methylation in several different loci in association with total IgE and childhood asthma[208-210]. In conclusion, significant differences in DNA methylation in the *NPSR1* promoter were detected in associations with asthma-related traits in adults and children, and in particular, with allergic asthma in children, which supports previous findings linking *NPSR1* to asthma and sensitization.

Smoking in adults showed rather large differences in methylation level compared with never-smokers. Current smoking was associated with 3.8% increase in DNA methylation at CpG site 1 (p = 0.005). In children, the same CpG site showed a negative association in relation to parental smoking in infancy (-0.52%, p = 0.02), although the absolute difference was small. At CpG site 8, both current and former smoking in the adult sample was associated with an increase in DNA methylation (8.74%, p = 0.042 and 6.36%, p = 0.013, respectively). Both CpG sites 1 and 8 reside in transcription factor (TF) binding sites for the TF Myb. In addition, a methylated C at CpG site 8 showed stronger TF binding in the EMSA, indicating increased transcription. Differential blood DNA methylation in association with active smoking has been observed by others. A systematic review of EWASs found 1,460 CpG sites to be associated with active smoking[211]. The most frequently reported sites were found in *AHRR*, *F2RL3* and *GRP15*. Moreover, several studies have found associations between tobacco smoke exposure *in utero* and differential DNA methylation[212-214]. In a meta-analysis of EWAS data from 13 cohorts, maternal smoking in pregnancy associated with nearly 6,000 differentially methylated CpGs. Persistence of DNA methylation patterns into later childhood was also detected[124]. In conclusion, several differentially methylated sites across the genome have been identified in relation to tobacco smoke exposure. In our study, smoking exposure was associated with differential methylation in the *NPSR1* promoter in both adults and children, with the largest differences for own smoking in adults, suggesting a link between environmental and genetic factors.

Traffic air pollution has been associated with differential DNA methylation[120, 215]. In study **II**, we did not observe any association between NO_x exposure at birth and current exposure at 8 years, in relation to DNA methylation in the *NPSR1* promoter in children from BAMSE. In study **III,** the suggestive genetic regions of importance in the association between traffic NO_2 and childhood asthma that were identified in the GWIS (*ADCY2*, *B4GALT5*, *DLG2* and *MOCOS*) were also investigated for differential DNA methylation in relation to traffic NO_2. NO_2 exposure at birth showed an inverse association with DNA methylation in one CpG site of *DLG2* (cg02275784, -2.7% per 10µg/m^3 NO_2 increase, $p = 1.21 \times 10^{-4}$) in blood from BAMSE's 8-year-olds. This CpG site is located in the gene body of *DLG2* at the position of a predicted enhancer element. Short-term diesel exposure in adults showed inverse associations of DNA methylation at eight CpG sites in the *DLG2* locus (-2% post- vs. pre-exposure for the CpG site with the lowest $p = 4.64 \times 10^{-5}$) and a positive association with two CpG sites (4% post vs. pre exposure for the CpG site with the lowest $p = 1.07 \times 10^{-4}$). In addition, decreased methylation was also identified at one *ADCY2* CpG site, and increased methylation was seen at one *MOCOS* CpG site.

Candidate gene studies have found differential methylation in relation to individual air pollution exposure in biosamples from blood[216-218], nasal airway epithelium[219], saliva[220] and in buccal cells[221, 222]. *In utero* exposure to air pollution has been associated with differential methylation in cord blood[123, 223, 224] and to global methylation in placenta[225]. Methylation changes within inflammatory genes have also been seen in mice[226]. A few EWASs have been performed in relation to air pollution exposure. A controlled experiment in adults showed associations between short-term diesel exhaust exposure and differential DNA methylation in > 2,800 CpGs in genes related to inflammation and oxidative stress[141]. In relation to long-term exposure, a recent meta-analysis in more than 1,500 newborns observed associations between prenatal NO_2 exposure and differences in DNA methylation in 3 CpGs in mitochondria-related genes[123]. In addition, differential methylation in antioxidant defense genes (*CAT* and *TPO*) was observed. A study using a crude comparison of children living in high vs. low pollution areas found differences in DNA methylation of more than 10% in 58 CpGs[227]. Overall, air pollution exposure has been associated with differential DNA methylation across the genome. In study **III**, we detected associations between air pollution exposure in children and differential DNA methylation in *DLG2*, which was confirmed in a controlled experiment of short-term exposure in adults.

In addition, in study **II,** higher methylation levels at *NPSR1* CpG site 5 were seen in children who had a blood sample drawn in July-September (1.38%, $p = 0.01$) as compared with those who gave blood in January-March. Birth month did not affect the methylation status. In a study of healthy individuals, hypo-methylation in long interspersed nucleotide element-1 (LINE-1) was seen in those who had a blood sample drawn in autumn and winter compared with in spring and summer, suggesting that seasonality should be taken into account in analyses of methylation levels[228]. In our analyses, adjustment for sample season did not affect the other results.

5.3 GENE EXPRESSION IN ASSOCIATION WITH GENETIC VARIANTS AND TRAFFIC AIR POLLUTION

In study **II**, transcription factors bound differently to CpG sites in the *NPSR1* promoter transcription start site depending on both genotype (rs2168890 at CpG site 2 and rs887020 at CpG site 9) and methylation status of the C allele (CpG site 2, 3 and 8). Since both disease status and environmental factors showed associations with differential methylation, this finding is of interest in terms of regulation of gene transcription.

In study **III**, the SNP with the lowest p value for interaction in the look-up evaluation was rs686237 (interaction p = 0.0016 in CHS). This SNP is located on chromosome 20, 40 kilobases (kb) upstream of *B4GALT5* and was identified as a lung eQTL of *B4GALT5* ($p = 1.18x10^{-17}$), where the C allele was associated with increased expression. eQTL analyses in blood from adults showed opposite direction where the C allele was associated with decreased expression ($p = 4x10^{-4}$). Further, in BAMSE's 16-year-olds NO_2 exposure at birth was associated with decreased expression of *B4GAL5* in blood, in homozygous carriers of the C allele, with a significant p value for genotype x NO_2 interaction (interaction p = 0.001) (**Figure 5.4**). The *ADCY2* SNPs were identified as eQTLs in blood from adults ($p = 4.50x10^{-4}$ for rs6886921), reference C allele being associated with increased expression. NO_2 exposure at birth was positively associated with gene expression of *ADCY2* in 16-yearsolds in BAMSE (p = 0.05), although no interaction between SNP (rs6886921) and NO_2 was detected. NO_2 exposure at birth was also positively associated with gene expression of *DLG2* in 16-year-olds (p = 0.0008).

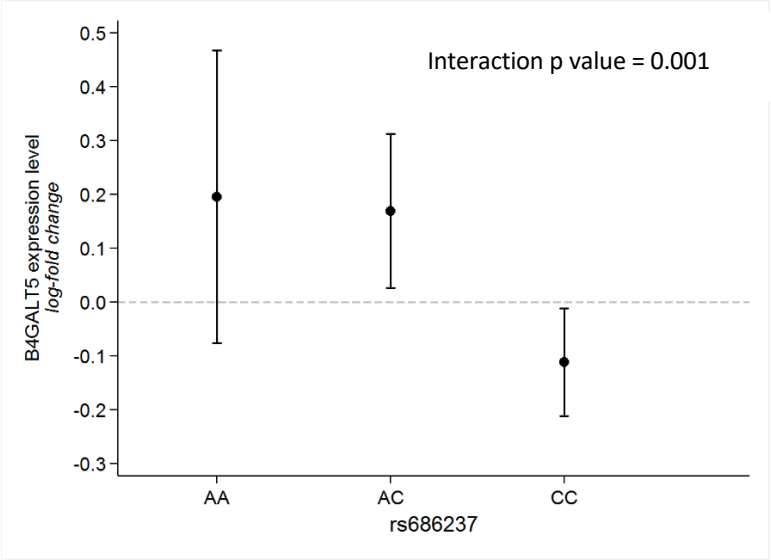

Figure 5.4 Adjusted coefficients and 95% confidence intervals for the association between NO_2 exposure at birth and peripheral blood gene expression levels of *B4GALT5* at 16 years of age in BAMSE, stratified by rs686237 genotype. The coefficient is the log-fold change in gene expression per 10µg/m^3 increase in NO_2 exposure.

Associations between air pollution exposure and gene expression profiles have been reported in animal models and in vitro experiments[229]. Few studies have been conducted in humans. In a comparison of children living in a highly polluted area vs. a less polluted area in the Czech Republic, differential genome-wide gene expression were detected in several genes, of which

C15orf16 and *DUSP15* showed the lowest p values for differential expression ($p = 4.97x10^{-7}$ and $2.59x10^{-6}$, respectively)[230]. Short-term exposure has been associated with changes in gene expression in genes related to inflammation, tissue growth and environmental defense mechanisms including oxidative stress[231-233] as well as protein degradation and coagulation[234]. To conclude, SNPs identified in the GWIS screening step showed significant eQTL effects in relevant tissues. Few studies on air pollution exposure in relation to gene expression have been performed, although evidence suggest that exposure is associated with differential expression of certain genes. In study **III**, the top hit replicating SNP in the *B4GALT5* locus provided supportive evidence for interaction with air pollution in relation to gene expression.

5.4 METHODOLOGICAL CONSIDERATIONS

This thesis is based on the well-characterized prospective birth cohort BAMSE as well as other similar birth cohorts with large number of participants, high retention rates within the cohorts, repeated assessments of health outcomes and exposures with reasonably good overlap in outcome and exposure definitions. A strength is the uniform outcome definitions obtained from detailed questionnaires repeatedly filled out during childhood, as well as objective measurements of sensitization. For a large proportion of children, SNPs were previously genotyped, using candidate gene and genome-wide approaches, and some of the children from BAMSE had DNA methylation and transcriptomic data available, which gave the possibility to study GxE at genetic, epigenetic and transcriptomic levels. Collaboration with other research groups, as well as the availability of comprehensive online resources, enabled functional follow-up analyses. However, both epidemiological and experimental studies are subject to limitations, related to study design or methods used, that should be taken into consideration when interpreting the results of these studies.

5.4.1 Random errors

In epidemiological studies, a sample of the population is studied. This sample will only give an estimate of the true population characteristics. Due to variability in the population, a randomly selected sample can introduce random errors. The precision will increase with increasing sample size. Confidence intervals are presented as an indicator of precision, where narrow confidence intervals represent higher precision. The precision of a study can also be affected by the prevalence of exposure and outcome or the quality of measurements. The precision of a study hence largely depends on random errors.

Studies I, II and III included subsamples of the original cohorts with adequate blood samples, informed consent and questionnaire data available, which to some extent limited the sample size. Studying GxE further reduced the sample size, since only a limited number of cases and controls had the combination of variables under study available. To enable observation of small effects of risk alleles an even larger sample size would have been preferred.

To increase the sample sizes, studies I and III were based on international consortia using meta-analysis and look-up approaches, resulting in analyses of 14,595 and 1,534 plus 1,602 children, respectively. The confidence intervals were rather narrow for study I, indicating good precision. In studies II and III, smaller sample sizes were available, which introduced some statistical uncertainty. Study IV included 2,359 children from the BAMSE cohort and the confidence intervals were rather narrow. However, the power for analysis of GxE in study IV was limited, due to the small sample size with genetic information available.

5.4.2 Systematic errors

Systematic errors, or bias, are errors that cannot be reduced by increasing the sample size. Systematic errors are related to the enrolment of study population, the assessment and definition of exposures and outcomes. Absence of systematic errors means high validity of a study. In epidemiology, systematic errors are usually divided into three categories: selection bias, information bias and confounding.

5.4.2.1 Selection bias

Selection bias will be present if the association between exposure and disease differs between subjects included and those not included in the study[127]. In prospective cohorts, subjects are typically enrolled before the outcome has occurred and therefore selection bias is unlikely.

Loss to follow-up may introduce selection bias. In BAMSE the participation rate has been high over time with a response rate of 84% for the questionnaire at 8 years and 78% at 16 years, with 60% and 63% participating in the clinical examinations at 8 and 16 years, respectively. The distribution of baseline characteristics, including parental allergy, were similar between the original cohort and the follow-ups, including the clinical examinations.

The inclusion criteria for each study are also subject to possible difficulties in obtaining a large enough and representative sample. The main analyses in studies I-III were dependent on which children participated in the clinical examinations, and gave blood and informed consent for genetic analyses. From the available samples, cases and a random selection of controls were selected for gene-specific and genome-wide genotyping.

5.4.2.2 Exposure assessment and outcome definitions in epidemiological studies

Information bias, or misclassification of exposure or outcome, can be either differential, meaning that the exposure or outcome depends on the other variable, or non-differential, if it does not depend on the other variable[128]. In BAMSE, differential misclassification of exposure is unlikely, thanks to the prospective study design where the exposure was assessed before the outcome occurred. However, parental heredity and allergic disease in early life might modify exposures later in life. Adjustment or restriction could control for that.

Mold, air pollution and food are complex mixtures of exposures where the exact disease-causative agents are not known. Assessment across cohorts may include analyses of heterogeneous variables, especially when combining data from different study areas and

countries. The choice of assessment method is critical and validation of the methods is important.

Visible mold will not necessarily be aerosolized and inhaled and there is a complex mixture of microbial agents and components in indoor environments[235]. In our study, exposure was assessed from parental responses to questionnaires, meaning that we could not elucidate the microbial content of the mold. Exposure assessment was performed in early life; however, bias might exist with over-reporting of exposure in children with symptoms of early wheeze or in families with parental allergic disease[236], which could lead to an overestimation of the associations between mold and outcomes in early life. In a recent study in BAMSE, risk estimates for asthma and rhinitis in relation to mold exposure were elevated among children both with and without parental allergic disease[172]. Also, over-reporting of dampness and mold in self-reported vs. inspectors reported observations has not been observed in relation to respiratory illnesses[37, 40].

Air pollution estimates have been modeled. In modelling, there is a fine balance between the incorporation of accessible predictors and adding too much uncertainty to the model. Validation studies have shown that both dispersion[146, 237, 238] and LUR estimations[239, 240] correlate well with outdoor measures. Outdoor levels modelled with dispersion modelling have shown good agreement with personal exposure measurement[241]. Good agreement was also seen between LUR and dispersion modelling for NO_2 estimates with a correlation of > 0.86, for areas included in study III[242].

Exposure to air pollutants was only estimated at residential and school addresses, whereas personal exposure and indoor exposure were not accounted for in the model. Since air pollution estimates were produced independently from the outcome assessments, any bias related to traffic air pollution exposure is assumed to be non-differential, possibly with attenuation of the true associations.

Smoking was assessed through questionnaires, with a possibility of underreporting. The association with methylation is assumed to be non-differential.

The FFQ used in study IV has not been validated in children, although the BAMSE study used a similar FFQ as a cohort of Swedish women. In this adult population, the validity and reproducibility of the FFQ-based TAC estimates have been assessed and a correlation of 0.27 (95% CI 0.06-0.46) was observed between lipophilic and whole plasma ORAC and FFQ-based ORAC estimates[243]. However, with respect to true exposure levels, TAC in plasma may not be the gold standard to estimate antioxidant intake, as both endogenous and exogenous factors besides dietary intake can affect TAC levels in plasma. FFQs may cause bias, since responders could find it hard to remember what and how much of each specified food item that was eaten during the previous 12 months. Further, individual eating habits could violate the estimation. Questionnaire data on 98 food and drink items were available in study IV. ORAC-TAC values from the US database were available for only 35 of these, although the major sources of antioxidant-rich foods were covered by both our FFQ and the

US database. The remaining 63 items in our FFQ with no TAC data available were dairy products, margarine or oils, meat dishes, fish, egg, pasta, rice, pizza, sweets or snacks, foods known to have low content of antioxidants in general. There might be differences in the antioxidant content related to different growing conditions and cooking methods of the foods consumed in our Swedish cohort in comparison to the foods in the US database, which could introduce errors in our TAC estimates. Still, we considered the TAC values to be valid in the BAMSE cohort of Swedish children and potential misclassifications are believed to be non-differential.

Misclassification of the outcomes is possible when the information is ascertained by parental questionnaires. Asthma or allergic symptoms might be reported more often among parents with allergic disease themselves due to a higher awareness of the disease. Early wheezing could be mistaken for airway infections and younger children have less distinctive asthma symptoms making it harder to diagnose true asthma. The disease definitions used in the meta-analyses have been harmonized and used in several international collaborations before[54, 106, 244]. In studies II and III, all children with a doctor's diagnosis of asthma ever, up to 8 years of age, were defined as cases. This broad definition enabled inclusion of many subjects, although with the potential limitation of a reduced contrast between cases and controls. For the asthmatics included in studies II and III, 49% had a current diagnosis at age 8 years, whereas 23% only had a diagnosis at 1 or 2 years of age in study II and, for study III, 26% were diagnosed at 1 or 2 years of age only. Sensitization was assessed using standardized methods. The IgE concentration in blood was measured for most cohorts, whereas ALSPAC and CAPPS used SPT to assess sensitization status.

5.4.2.3 Confounding in epidemiological studies

A confounder must be associated with both the disease and the exposure, and must not be an effect of the exposure[128]. If a confounder is not controlled for in the analysis, an observed association may in fact be due to the confounding factor. Confounding can cause either underestimation or overestimation of effects. It could also be strong enough to reverse the apparent direction of effect[128]. In my studies, adjustments were made for known or suspected confounders for asthma and allergic disease as well as for variables influencing the effect estimates.

Dampness and mold may be particularly prevalent in poorly maintained housing and may additionally contribute to poor health in populations who are already living with an increased burden of disease[34]. However, socioeconomic status was not a confounder in the association between mold and asthma or allergic disease and was therefore not adjusted for in study I[173]. Since the fungal profile could differ between different homes and different study areas, residual confounding cannot be excluded for this exposure. For most cities, low socioeconomic status is related to living in highly polluted areas. In Stockholm, the opposite is true. Children from families with high socioeconomic status tend to live in the inner city, where the levels of air pollution are highest. Therefore, all analyses of traffic air pollution estimates were adjusted for municipality.

We did not have information about maternal diet in pregnancy or early life dietary TAC estimates (study IV). Further, any potential bias related to an unhealthy diet cannot be ruled out, since a high dietary TAC might just reflect the relative intake of protective and harmful foods that influence the pro-oxidant-antioxidant balance[245]. However, adjustment for meat intake did not affect the obtained estimates.

Disease-related modification of consumption is a methodological challenge when studying diet, mainly in cross-sectional studies, but also to some extent in longitudinal studies. Children who reported allergic symptoms related to fruits and vegetables, and/or who avoided these because of allergic symptoms, were significantly more represented in the lowest tertile of TAC intake. After exclusion of children who avoided fruits or vegetables due to allergic symptoms, the results remained comparable for the association between dietary TAC and sensitization to inhalant allergens. Since TAC aims to reflect the whole diet, it may be less affected by disease-related modification of consumption affecting single food items, as reported by Rosenlund et.al[73]. Indeed the sensitivity analysis showed no evidence of disease-related modification.

Residual confounding cannot be ruled out because of unmeasured confounding or measurement errors in confounders already adjusted for and this applies to all studies in this thesis.

5.4.2.4 Assessment of genetic, epigenetic and transcriptomic data

Genotyping errors cannot be excluded, although these are believed to be non-differential. Whole genome microarrays may in fact not cover the whole genome. Despite of imputation (from HapMap2 data) of SNPs in the genome-wide analyses, which resulted in about 2.5 million SNPs, a denser genomic coverage, through 1000 Genomes imputation might have reduced the risk of false negative findings in study III. At the time when our GWIS study was initiated, only HapMap2-imputed data were available for all cohorts. However, 1000 Genomes data was available for BAMSE and a comparison between HapMap2 and 1000 Genomes results for two of the top hits (*B4GALT5* and *ADCY2*) did not identify any other SNP that was not already in high LD with the already identified SNPs. Only common variants, with a minor allele frequency > 5% were investigated. To cover rare variants next-generation sequencing technologies would be preferred[246].

We observed rather small absolute differences in methylation levels in relation to exposures and outcomes investigated (studies II and III). It should be noted that methylation differences may involve technical variation related to the bisulfite treatment, EpiTYPER or Illumina 450K microarray chip, and adjustment for batch effects is therefore often applied. Unsuccessful conversion of an unmethylated cytosine to a thymine may cause false positive findings[247], which should be random with respect to the exposure of interest, leading to non-differential misclassification. The Illumina 450K microarray covers CpG sites distributed over the genome, including CpG islands, CpG island shores, shelves, promoters, gene bodies and intergenic regions; still it only includes about 2% of the total number of CpG sites. In

study III, functional follow-up of CpG sites 50 kb up- or downstream of the identified genes were investigated in relation to air pollution exposures and regions of 250 kb up- or downstream of the SNP were investigated for gene expression. There is no consensus on the optimal distance for these types of analyses, though strongest associations between SNPs and CpG sites have been detected for SNPs located close to the CpG sites, within 500kb[248]. Regarding genome cover in the transcriptomics analyses, the Affymetix HTA 2.0 microarray includes both coding and non-coding transcripts covering all known genes and transcript isoforms.

5.4.2.5 Confounding in genetic studies

Genotypes are known to differ between ethnic groups, and ethnicity was adjusted for in analyses of genotypes by restriction (to e.g. Caucasians) or adjustment. A SNP can affect methylation, if it coincides with a CpG site or by being a methQTL. In study II, analyses were restricted to homozygous carriers of a SNP creating a CpG site. In study III, CpG site SNPs were excluded. For both studies II and III, methylation and transcriptomic samples were randomized into batches and batch effect was adjusted for.

Tissue specificity and cell type adjustment

Methylation and gene expression patterns can be tissue-specific. To look for differential methylation and gene expression in relation to respiratory and allergic diseases, lung and blood tissue would be of relevance to study. Furthermore, specific types of lung tissue, e.g., epithelial or alveolar cells, might show different direction of effects. Bronchial epithelium is the first contact with cellular responses for environmental exposures such as air pollution. Samples are most commonly extracted from peripheral blood for reasons of convenience, and its relevance is supported by the systemic nature of allergy and to some extent, asthma.

Without cell type adjustment, confounding by, e.g., difference in cell type composition between cases and controls can be present. Cell separation or adjustment for cell type is therefore of importance in the assessment of DNA methylation and gene expression. In study II, no cell type correction was performed, which potentially could have biased our results. However, whole blood and cell-sorted blood from six healthy males suggested constant methylation levels across cell types in the *NPSR1* gene, indicating that whole blood was a reliable source for DNA methylation analysis. In study III, cell type correction was done using the Houseman algorithm[167] in combination with reference data from the aforementioned cell-sorted blood[117]. Measured cell counts were adjusted for in the transcriptomic data.

Stability over time

DNA methylation and gene transcription are dynamic processes influenced by numerous factors such as genetic variants, age, exposure and health conditions[125]. Assessed levels may depend on timing of sampling (as seen in study II, where season of blood sampling was associated with differential DNA methylation) and stability of DNA methylations may

depend on the genomic locations of the CpG sites. At a population level, maternal smoking during pregnancy as well as air pollution exposure during pregnancy have been associated with differential methylation in cord blood with persistence of DNA methylation patterns into later childhood[123, 124]. However, since DNA methylation and gene expression was assessed at one point in time (studies II and III), these analyses should be considered to be cross-sectional. The assessed levels may not represent long-term effects in our studied material and the direction of effect cannot be ruled out.

Even if we can assume that the associations seen in our studies are true, and not explained by residual or unmeasured confounding, it may still not be possible to draw conclusions about causality. The identified disease-associated SNP might not be the SNP that is truly associated with disease. It can be in LD with a SNP in another locus that causes the effect and there might be several mediating steps in between, e.g., DNA methylation, gene expression, protein expression or further interactions with other loci. Therefore, the identified disease-associated loci from the GWIS were further investigated in functional follow-ups.

5.4.3 Generalizability of results

The birth cohorts included in studies I-III are population-based. For BAMSE, 75% of those eligible were included in the final cohort. Parental smoking was more common among non-participants, whereas allergic heredity and other factors did not have an impact on participation. The cohort was therefore considered reasonably representative for the general population of the same age, although not fully representative regarding smoking exposure[129].

The objective of studies I-III was to gain knowledge in the underlying mechanisms of allergic disease. If equal underlying biological mechanisms can be assumed, the observed associations could be generalizable to most populations of European and non-Hispanic white origin.

Our results in study IV could be considered generalizable to a Swedish population of children, since the study population represented the original population-based BAMSE cohort, with respect to baseline characteristics. Generalizability to most populations could be assumed, given the belief of equal underlying mechanisms.

5.5 STUDYING INTERACTIONS BETWEEN GENES AND ENVIRONMENT

5.5.1 Statistical models

Different approaches were used to study interactions in this thesis. Traditional models for evaluating interaction were applied in studies I, III and IV by investigating statistical interaction on a multiplicative scale. To enable interpretation of statistically significant interactions (often presented as a significant p value for interaction), stratifications are commonly performed. This was done in studies I, III and IV to visualize the strength of association and the precision (by presentation of confidence intervals). In study I, the minor

allele frequency (MAF) was 12%, whereas MAF ranged between 9% and 35% for the top hit SNPs in study III. Dominant coding of SNPs was tested in all studies in order to increase power.

Traditional interaction models have been shown to be appropriate for targeted and hypothesis-driven analyses. Assessment of interactions in a hypothesis-free manner, as in GWIS, demands very large sample size and power easily becomes a concern. Different GWIS approaches have been subject to much attention in the past years[249-252] given that GxE effects have been very challenging to identify.

In our GWIS (study III), a sufficient sample size was definitely a concern. Power improvement was addressed by performing meta-analysis, and using two alternative statistical approaches, the 2 df test[112] and the two-step approach[113]. The 2 df test was constructed to detect main genetic effect in either exposed or unexposed subjects and to detect main effects while taking the environmental exposure fully into account, although it was not optimized to primarily detect interaction effects. The two-step approach (see section 4.7.1) was used in order to identify loci of relevance for NO_2 interaction, utilizing a low significance threshold with no penalty of multiple testing correction. The much lower number of SNPs that passed the first step was then evaluated for interaction using logistic regression. Since fewer SNPs had to be adjusted for multiple testing, the power was increased. Even if we found genome-wide interaction effects in the discovery dataset using the two-step approach (*MAGI1* SNPs), these SNPs did not replicate in CHS or CAPPS/SAGE. In our case, the standard interaction model provided candidates (*ADCY2*, *B4GALT5*, and *DLG2*) that were eventually supported by functional analyses.

While the application of new methods to detect gene-environment effects (e.g., the 2 df test or the two-step approach) is one way to overcome power challenges, we believe that gene-environment effects should be evaluated also at the epigenetic and transcriptomic levels to functionally support GWIS top hit findings.

5.5.2 The role of functional studies and bioinformatics resources

Upon identification of GxE effects, a next step would be to understand the functional relevance of this interaction. Identified SNPs are often located outside the protein-coding regions operating as regulatory SNPs with unknown functional consequences[253], and are usually part of a larger region of correlated variants (in strong LD). Both SNPs and environmental exposures could affect methylation and gene expression, and DNA methylation could regulate gene expression[254]. Therefore, a GWIS can be seen as a screening tool to identify genetic loci that may interact with the other variables under study.

DNA methylation has been suggested to be involved in mediating GxE effects, and several examples exist in the literature related to air pollution exposure[120, 215]. In studies II and III, we investigated whether environmental exposures influenced the methylation levels of CpG sites in the asthma candidate gene *NPSR1* and CpG sites in asthma-associated loci identified in the GWIS. We found that smoking and air pollution exposure were both associated with

differential methylation in peripheral blood cells of the investigated regions, indicating a link between environmental and genetic factors. Further, study III investigated suggestive GxE effects at the transcriptomic level. SNP main effect on gene expression was observed as well as association between NO_2 exposure on gene expression, with significant interaction for genotype.

Thanks to public repositories and available datasets through different collaborations, look-up of research findings was possible in different tissues, allowing for DNA methylation and gene expression analyses in study III. Further, GxE could also influence the expression of proteins encoded by the genetic variant. Protein expression data was not available in the BAMSE material obstructing analyses of differential expression. However, results from the Human Protein Atlas showed that the identified genes were expressed at protein levels in normal respiratory tissue and smooth muscle tissue[145].

5.6 REPLICATION OF GENE-ENVIRONMENT INTERACTIONS

Despite a large sample size in study I, interaction with the *GSTP1* genotype could neither be confirmed nor ruled out. This highlights the complexity of studying interactions involving genes in a multifaceted oxidative stress defense system. Several studies on air pollution in relation to childhood asthma[81, 83, 84, 255] as well as lung function[85] have shown evidence for interaction with *GSTP1* and *TNF* genotypes, although with different direction of effects. These candidate genes were initially identified as having a main effect on asthma and allergic disease, which might explain observed heterogeneity in interaction analyses.

In our GWIS, a meta-analysis was performed including, at that point in time, all available cohorts with information on childhood asthma, NO_2 exposure estimates and GWAS data; totaling 454 asthmatics and 1,080 never-asthmatics. Although promising loci were identified, we did not reach genome-wide significance for the interaction analyses. Likewise, many previous GWIS efforts to detect GxE effects for respiratory disease and lung function have failed in reaching genome-significant interaction p values, in studies with a sample size of 20,000[187] or even up to 50,000 subjects[186]. Thus, new loci have been challenging to discover using standard interaction analyses with stringent genome-wide significance demands, even in well-powered studies. In addition, well-studied candidate genes for GxE have been difficult to replicate using GWAS data[185].

Research in large-scale consortia demands good quality data with harmonization of variables over participating study centers. The difficulties in replicating findings may relate to different environments (with respect to exposure levels or constituents) or populations (including different allele frequencies). The effect of genotype and environment might influence the phenotype differently in different populations or environments, which could be related to co-exposures or unmeasured confounding. Meta-analysis may therefore not necessarily increase the power because of heterogeneity in study design. However, allele imputation, adjustment

for population stratification and using harmonized exposure and outcome definitions can meet some of these uncertainties.

In study I, the prevalence of mold exposure varied between 27% and 68% between cohorts, which could possibly be explained by different climate or housing conditions, a heterogeneity that may have affected the power to identify interaction. Similarly, in study III, the range of estimated air pollution levels varied between cohorts in the discovery meta-analysis, where BAMSE had lower exposure levels than the other two cohorts (**Figure 5.5**).

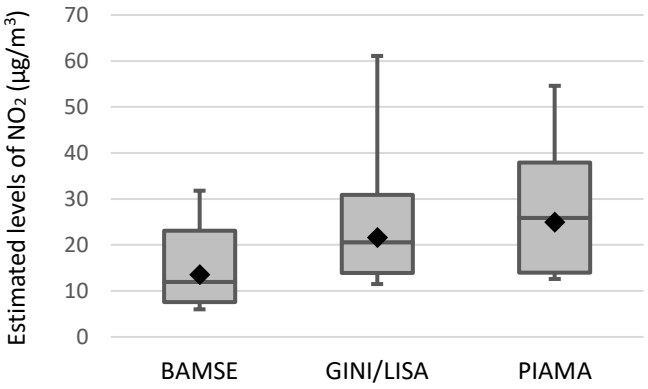

Figure 5.5 Range of NO_2 exposure levels (min-max) for the cohorts included in the discovery meta-analysis. Boxes represents 5^{th} to 95^{th} percentiles and median. Black diamonds indicates mean exposure levels.

Sensitivity analysis with exclusion of BAMSE from the meta-analysis showed same magnitude and direction of interaction beta values for BAMSE as for meta-analysis of GINI/LISA and PIAMA cohorts. This indicated that the difference in exposure levels did not influence the discovery results to any major extent. Stringent meta-analysis of the top hit SNPs between the European and North American cohorts was not performed due to heterogeneity in study designs. Air pollution data from CHS was assessed using a different method and the exposure was obtained for when children were 8 years of age.

SNPs that have shown genome-wide significance for main effects in GWASs, such as the *ORMDL3* and *GSDMB* loci, have been analyzed in many studies and for different ethnicities, with results indicating true associations with asthma. SNPs with a lower effect estimate may be difficult to identify in low-powered studies. In genome-wide studies, multiple testing correction is necessary to avoid false positive findings, with the penalty of increasing the likelihood of false negatives. Less significant SNPs from GWASs may still be of importance in the disease mechanism and should not be discarded from further study. Our approach was therefore to perform functional analyses of all the top GWIS findings that also showed significant interaction in the look-up data set, to further look for evidence of their involvement in childhood asthma at a more detailed level.

New genetic findings were revealed in the GWIS when we looked at interactions followed by methylation and gene expression associations, which were verified in different tissues. Consistent epigenetic effects surviving multiple test corrections in adults and children for one

of the genes (*DLG2*) using both experimental and epidemiologic data, speak for a solid finding.

It should be noted that the SNP with the lowest p value in the look-up evaluation, rs686237 on chromosome 20, showed a rather complex interplay with exposure and asthma risk. First, the direction of interaction effect differed between the discovery meta-analysis and the primary look-up dataset. We observed that AC/AA carriers had the highest risk of asthma following NO_2 exposure in the discovery datasets (OR 1.69 95% CI 1.08-2.64), whereas CC carriers had the highest risk in CHS (OR 1.21 95% CI 1.04-1.41). Detailed analyses in CHS revealed strong confounding effects in the data set, and the differences in asthma risk might be influenced by residual confounding. Rs686237 was found to be a strong eQTL for expression of *B4GALT5* in the lung and was further identified as an eQTL in whole blood. In lung tissue, the C allele was strongly associated with increased expression (risk allele in CHS), but in blood (GTEx), the opposite pattern was observed, with the A allele being clearly associated with increased expression (risk allele in the discovery datasets). Stratified analyses by genotype in BAMSE (blood) confirmed the GTEx findings in that allele A carriers displayed increased *B4GALT5* expression following NO_2 exposure, but allele C carriers displayed decreased expression.

In summary, the GWIS study support the notion that GxE effects are important in the pathogenesis of asthma. Although our findings showed a complex interplay, the study highlights the importance of studying GxE at different levels by analyzing genetic variations, gene regulation and gene expression to understand disease pathophysiology. Since the role of these genes and genotypes in the pathogenesis of asthma is not known, further studies are warranted.

6 CONCLUSIONS

Based on the studies in this thesis, the following conclusions can be drawn:

I: Early mold exposure was associated with early wheezing and nasal symptoms in children. Modification by a functional genotype in the *GSTP1* gene, involved in the anti-oxidative stress system, could not be confirmed.

II: *NPSR1*, a gene identified for susceptibility to asthma, showed differential DNA methylation in the promoter region in relation to adult and childhood asthma, in particular childhood allergic asthma. Environmental exposures, such as smoking and season of blood sampling, were also associated with differences in DNA methylation levels.

III: Using a genome-wide approach we found that interactions between single nucleotide polymorphisms and traffic air pollution exposure are likely to be important for asthma development. DNA methylation and gene expression analyses provided supportive evidence for interaction with air pollution for *ADCY2*, *B4GALT5* and *DLG2*.

IV: An antioxidant-rich diet in childhood may protect against development of allergic sensitization to inhalant allergens. Our results indicated that the beneficial effect was predominant in conditions of low oxidative stress, e.g., exposure to low levels of air pollution.

7 FUTURE PERSPECTIVES

Evidence today suggests that a combination of genes and environment plays a crucial role in asthma and allergic disease development. The results in this thesis imply that genetic variants could influence associations between environmental exposures and asthma and allergic disease. Further, interaction between environmental factors was indicated in relation to disease development.

Studying GxE interactions in asthma and allergic disease is a challenge with respect to study power, especially at the genome-wide level. Adequate sample size, high quality data with detailed genotyping and exposure assessments, as well as deep phenotyping (possibly with the help of objective biomarkers) are important factors to increase study power. The use of international collaborations and consortium data can contribute with enlarged sample sizes to investigate different sub-phenotypes with sufficient power, although harmonization of data is a delicate issue and it is important to put efforts into harmonizing the variables between studies. Collaborations and publicly available databases, which have already become commonly used in genetic studies, can provide opportunities for look-up of research findings in other settings.

There is a need for improved statistical methods and ways to follow up GxE findings with functional studies. Future research should focus on studying more complex interaction patterns, such as interactions between several genetic variants at once or analyzing haplotype structures. The integration of different omics data raises a need for method development on how to handle the amount and complexity of information obtained from simultaneous analysis of genome-wide genetic-, epigenetic-, transcriptomics-, proteomics- and metabolomics- data.

The results in this thesis indicated a possible link between environmental exposures and epigenetic as well as transcriptomic events, which should be further investigated in longitudinal set ups. Relevant tissues and cell types (e.g., airway epithelium) should be studied in relation to methylation and gene expression in a longitudinal manner, in order to get one step closer to deciphering causality and investigating the stability of changes over time. In particular, looking at changes over time in relation to exposure, or before and after disease onset, in the same individual could provide new insights into disease mechanisms. Through increased understanding of these patterns, there is potential for developing new diagnostic tools.

Pregnancy cohorts that can consider *in utero* exposures would be preferred, or multi-generation studies, with longitudinal bio-sampling of blood and other relevant tissues. Follow-up of these individuals up to adulthood would enable investigation of early risk factors for later development of asthma related lung diseases such as COPD.

8 POPULÄRVETENSKAPLIG SAMMANFATTNING

Astma, hösnuva och eksem är exempel på allergisjukdomar. Förekomsten av allergisjukdomar hos barn är stor. Många barn i Sverige kommer att ha någon form av allergisjukdom under sin uppväxt och astma är den vanligaste kroniska sjukdomen hos barn. Den största riskfaktorn för allergisjukdom är ärftlighet, men under de senaste decennierna har sjukdomsförekomsten ökat, vilket tros beror på förändringar i den omgivande miljön. Föräldrars rökning, mögel i bostaden, trafikrelaterade luftföroreningar, kost och kontakten med mikroorganismer är exempel på miljöfaktorer som har kopplats samman med allergisjukdomar hos barn. I dagsläget vet man inte exakt vilka mekanismer som ligger bakom sjukdomsutvecklingen men arv och miljö tycks samspela.

Syftet med den här avhandlingen var att studera genetiska faktorers inverkan på sambanden mellan miljöfaktorer och allergisjukdom. Avhandlingen omfattar fyra delstudier vilka alla inkluderar epidemiologiska data från den svenska födelsekohorten BAMSE, en studie på 4000 barn i Stockholmsområdet som följts upp till 16 års ålder. Delarbete I och III inkluderar också data från liknande födelsekohorter i Europa och Nordamerika. Delarbete II innefattar, förutom BAMSE, en kohort på vuxna med kronisk luftvägssjukdom. Information från experimentella studier och från publika databaser har också nyttjats.

I **delarbete I** undersöktes sambanden mellan tidig exponering för fukt och mögel i bostaden och allergisjukdomar upp till skolåldern, samt inverkan av genen *GSTP1*. *GSTP1* kodar för ett enzym som finns i lungan och som kan hjälpa till i kroppens försvar mot toxiska substanser som inandats. Naturligt förekommande varianter av *GSTP1* kan göra vissa individer mer känsliga än andra för exponering av inandningsbara substanser. Studien inkluderade totalt 14 495 barn, och innefattade BAMSE-studien samt fem andra födelsekohorter. Exponering för fukt och mögel var relaterat till tidig väsande andning (”wheeze”) samt hösnuva upp till skolåldern. Den ökade risken för hösnuva sågs främst hos barn som hade en viss variant av *GSTP1*. Sammanfattningsvis indikerade resultaten att exponering för fukt och mögel i hemmet under de två första levnadsåren var kopplat till allergisjukdom hos barn, en påverkan som främst sågs hos barn som hade en viss genetisk variation.

I **delarbete II** studerades den tidigare identifierade astmagenen *NPSR1*. Epigenetiska förändringar, det vill säga kemiska förändringar av vår arvsmassa som kan påverka genernas uttryck och funktion, har föreslagits vara kopplingen mellan miljöexponering och sjukdomar som astma. I detta delarbete undersöktes om miljöexponeringar och astma var kopplade till sådana epigenetiska förändringar i *NPSR1*. I studien sågs samband mellan epigenetiska förändringar i relation till astma hos både vuxna och barn, och speciellt i relation till allergisk astma hos barn. Hos vuxna rökare identifierades också epigenetiska förändringar, vilka inte sågs i samma utsträckning hos icke-rökare. Epigenetiska skillnader sågs också beroende på vilken tid på året blodprovet togs. Resultaten från denna studie tyder på att det finns samband mellan miljöexponering och epigenetiska förändringar, vilka i sin tur föreslås påverka reglering av geners uttryck.

Syftet med **delarbete III** var att studera hela arvsmassan för att identifiera barns känslighet för trafikrelaterade luftföroreningar och risken att utveckla astma. Studien inkluderade mer än 3 300 barn från tre Europeiska och två Nordamerikanska kohorter. Exponering för förhöjda halter av luftföroreningar under första levnadsåret var kopplat till astma, men bara hos barn som hade ett visst genetiskt arv. Två av de gener som identifierades, *ADCY2* och *DLG2*, har i tidigare studier kopplats till lungfunktion och lungsjukdomen KOL. Uppföljande studier gjordes för att studera genreglering och genaktivering i blod och lungvävnad. Luftföroreningsexponering kunde kopplas till epigenetiska förändringar i *DLG2*, vilket studerades hos barn från BAMSE-studien samt hos vuxna individer som utsatts för dieselavgaser i exponeringskammare under kontrollerade former. En av de identifierade genetiska variationerna (rs686237) visade sig påverka uttrycket av en närliggande gen, *B4GALT5*. Detta testades i lungvävnad hos vuxna individer. Hos barnen i BAMSE-studien sågs också en koppling mellan luftföroreningsexponering och uttryck av *B4GALT5* i blod, och detta samband påverkades i sin tur av vilken variant av rs686237 som barnet hade. Sammantaget visade resultaten att barn med ett visst genetiskt arv påverkar deras känslighet för luftföroreningar och utveckling av astma. Studien lyfter fram betydelsen av att studera arv och miljö på flera nivåer genom att analysera genvariationer, reglering och genuttryck för att identifiera sjukdomsmekanismer.

I **delarbete IV** undersöktes huruvida en kost rik på antioxidanter i 8-års åldern påverkade nyinsjuknandet i astma, hösnuva eller sensibilisering mot luftburna allergen vid 16 års ålder. Flera av barnen undvek intag av vissa frukter eller grönsaker på grund av allergiska symptom mot dessa, vilket kan påverka de studerade sambanden mellan kostens potentiellt skyddande effekt i förhållande till allergisjukdom. Tidigare studier som undersökt effekterna av ett högt antioxidantintag har visat tvetydiga resultat. Många studier har heller inte tagit hänsyn till denna sjukdomsrelaterade påverkan på kostintaget. Resultaten från delarbete IV visade att de barn som hade ett högt intag av antioxidanter via kosten hade en minskad risk för sensibilisering. Effekten kvarstod efter att vi tog hänsyn till sjukdomsrelaterad påverkan av kostintaget. Vidare kunde den gynnsamma effekten av antioxidanter bara identifieras hos de barn som levde i områden med lägre halter av luftföroreningar. Genetiska variationer i två gener kopplade till antioxidativa mekanismer (*GSTP1*) och inflammation (*TNF*) påverkade inte de studerade associationerna.

Sammanfattningsvis tyder resultaten i denna avhandling på att genetiska variationer påverkar sambanden mellan miljöexponeringar, såsom fukt och mögel samt luftföroreningar, och astma och allergisjukdomar hos barn. Samverkan mellan miljöfaktorer verkar också kunna påverka utvecklingen av allergisjukdom. Det verkar också finnas samband mellan miljöexponeringar och epigenetiska förändringar såväl som geners uttryck.

9 ACKNOWLEDGEMENTS

Many people have contributed to making this thesis an exciting and joyful journey and I would like to thank all of you who have helped me along the way.

First of all I would like to thank the BAMSE participants, their parents, and the participants in all the other cohorts and studies included in this thesis, for giving their time and effort.

Erik Melén, my main supervisor who is always so positive and full of ideas. Thank you for all the positive energy you bring to each project and for your encouragement. I am grateful to have had you as my supervisor on this adventure of genetic epidemiology. You have showed me that hard work and a positive attitude open doors to possibilities and great collaborations. 'Cause nothing's impossible, when we're a team.

Anna Bergström, my very kind co-supervisor. You have always kept your door open to me and have taken the time to sit down and discuss everything from research to family life. Thank you for all of the good advice you have given me and for showing me that non-genetic epidemiology is also fun. The precision and thoroughness with which you approach everything you do is truly inspiring.

Göran Pershagen, my co-supervisor and the one who welcomed me to this fantastic group of people. You impress with your canny expertise – your critique and critical questioning is always spot on.

Tove Rylander Rudqvist, my mentor. Thank you for the nourishing lunch talks and for giving your time and sound advice.

Anders Ahlbom and **Ulla Stenius**, former and present chairmen of the Institute of Environmental Medicine, for providing the possibility to perform research in an excellent research environment.

I would like to thank all the co-authors for sharing your time and invaluable knowledge. It has been truly inspiring to work with so many different research groups from different disciplines. **Christina Tischer** and **Joachim Heinrich** for allowing me to take part in your interesting project. Thank you **Lovisa Reinius**, **Nathalie Acevedo**, **Cilla Söderhäll**, **Annika Scheynius** and **Juha Kere** for a cohesive collaboration in the *NPSR1* project and other studies, and Cilla and Juha, for your contributions to the GWIS project. All of the researchers involved in TAG, for your enthusiastic engagement, productive meetings, and positive feedback along the way. A special thanks goes to **Chris Carlsten** and **Michael Brauer** who have been a devoted part of the GWIS project from day one who have always taken the time to give input on the project's progress. **Ashish Kumar** and **Annika Sääf** for all of the hours spent on the GWAS dataset, and Annika for your work with the *NPSR1* project. **Tom Bellander**, for sharing your extensive knowledge of air pollution. **Matteo Bottai**, for your statistical expertise in the GWIS project. **Alicja Wolk**, **Susanne Rautiainen**, and **Niclas**

Håkansson, thank you for introducing me to nutritional epidemiology, teaching me the concept of TAC and providing me with data.

The BAMSE team. The always smiling and enthusiastic **Inger Kull** and **Magnus Wickman**. **Sara Nilsson**, **Eva Hallner** and **André Lauber** for your much appreciated work at the BAMSE secretariat. The very helpful statisticians **Tomas Lind** and **Niklas Andersson**. Niklas, thank you for being such a lovely roommate and for your patience whenever I needed to voice my statistical and technical issues.

Colleagues at the unit of Environmental Epidemiology, fellow PhD students and post-docs. **Olena Gruzieva**, my former wonderful roommate. Thanks for all the ways you have supported me, from complicated research issues to taking coffee breaks. **Simon Kebede Merid**, my work would have taken another four years without the help of youR language skills. **Jessica Magnusson**, for the good times during and after work hours and for your much appreciated support during my kappa writing. **Erica Schultz** for your constant supply of positive energy. **Sandra Ekström** and **Alva Wallas**, thanks for keeping me in good shape. **Jesse Thacher**, **Andrei Pyko**, **Auriba Raza**, **Åsa Persson** for your great company. To the veterans **Charlotta Eriksson** and **Michal Korek**, high five! **Jenny Hallberg**, **Jennifer Protudjer**, **Sara Fagerstedt**, **Elaine Fuertes**, **Elin Dahlén**, **Emma Johansson**, **Maria Ödling**, **Petter Ljungman**, **Marcus Dahlquist**, **Per Thunqvist**, **Petra Um Bergström**, **Marina Jonsson**, **Åsa Neuman**, **Marit Westman**, **Natalia Ballardini**, **Mirja Vetander**, **Björn Nordlund**, **Anna Undeman Asarnoj**, **Helen Rosenlund**, **Johanna Penell**, **Sviatlana Panasevich**, **Ulrika Hellberg**, **Susanne Lundin**, **Mare Löhmus Sundström**, **Antonis Georgellis** and **Anne-Sophie Merritt**, it has been a privilege to work with all of you.

A thanks to the members of all the journal clubs I have attended, for the challenging and interesting discussions.

Sist men inte minst vill jag tacka alla er som finns i mitt liv varje dag och som fyller mig med positiv energi: **Michelle**, du är bäst! Tack för att jag får ta din vänskap för givet. **Elin**, för att du är en bra förebild som får mig att våga mer. Tack för att du svängde förbi KI en kortis. Tack för alla skratt och goda middagar mina kära biomedicinvänner, aka T.O.M. **Fanny**, **Hanna**, **Mariann**, **Niosha**, **Sofia W.L** och **Sofia A.W**. Alla små och stora i **Lillens Sällskap**, bästa grannarna ever. **Sofi** och **Jo**, tack för arbetsron. Alla fina **klättrare** (och ni andra som hör därtill) som delar mitt intresse för problemlösning. **Mamma** och **pappa**, för att ni är så kloka och okomplicerade. Tack för bra gener och en utmärkt uppväxtmiljö. **Staffan**, för att jag fått äran att vara lillasyster till den snällaste storebror man kan önska sig. Kära **Anne**, tack för att du tar så bra hand om oss allihop! Jämt. Oavsett. **Bosse**, tack för att du alltid varit mån om att jag ska ha ett bra arbetsklimat. **Minna** och **Hannes**, tack för allt mysigt häng och alla roliga äventyr. Jag ser fram emot den långa fortsättningen.

Andreas, du vet hur mycket du betyder för mig. Tack för att du inspirerar mig till att tänka lite djupare och klättra lite högre. **Teodor** och **Mira**, de finaste personerna i världen. Jag älskar er, mest av allt.

10 REFERENCES

1. Lichtenstein P, Svartengren M. Genes, environments, and sex: factors of importance in atopic diseases in 7-9-year-old Swedish twins. Allergy. 1997;52(11):1079-86.
2. Thomsen SF, Ulrik CS, Kyvik KO, von Bornemann Hjelmborg J, Skadhauge LR, Steffensen I, et al. Genetic and environmental contributions to hay fever among young adult twins. Respir Med. 2006;100(12):2177-82.
3. Asher MI, Montefort S, Bjorksten B, Lai CK, Strachan DP, Weiland SK, et al. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. Lancet. 2006;368(9537):733-43.
4. World Allergy Organization. White Book on Allergy [internet]. Milwaukee, Wisconsin: World Allergy Organization; 2011. [cited 2017 Mar 21]. Available from: http://www.worldallergy.org/UserFiles/file/WAO-White-Book-on-Allergy_web.pdf.
5. World Health Organization. Global surveillance, prevention and control of chronic respiratory diseases: a comprehensive approach [internet]. Geneva: World Health Organization; 2007. [cited 2017 Mars 21]. Available from: http://www.who.int/gard/publications/GARD%20Book%202007.pdf.
6. Murphy K, Travers P, Walport M. Janeway's immunobiology. 7th ed. New York and London: Garland Science, Taylor and Francis Group; 2008.
7. Galli SJ, Tsai M, Piliponsky AM. The development of allergic inflammation. Nature. 2008;454(7203):445-54.
8. Wickman M, Asarnoj A, Tillander H, Andersson N, Bergstrom A, Kull I, et al. Childhood-to-adolescence evolution of IgE antibodies to pollens and plant foods in the BAMSE cohort. J Allergy Clin Immunol. 2014;133(2):580-2.
9. Bunne J, Moberg H, Hedman L, Andersson M, Bjerg A, Lundback B, et al. Increase in Allergic Sensitization in Schoolchildren: Two Cohorts Compared 10 Years Apart. J Allergy Clin Immunol Pract. 2017;5(2):457-63 e1.
10. Martinez FD, Vercelli D. Asthma. Lancet. 2013;382(9901):1360-72.
11. Roberts G, Xatzipsalti M, Borrego LM, Custovic A, Halken S, Hellings PW, et al. Paediatric rhinitis: position paper of the European Academy of Allergy and Clinical Immunology. Allergy. 2013;68(9):1102-16.
12. Dozor AJ. The role of oxidative stress in the pathogenesis and treatment of asthma. Ann N Y Acad Sci. 2010;1203:133-7.
13. Mak JC, Chan-Yeung MM. Reactive oxidant species in asthma. Curr Opin Pulm Med. 2006;12(1):7-11.
14. Erzurum SC. New Insights in Oxidant Biology in Asthma. Ann Am Thorac Soc. 2016;13 Suppl 1:S35-9.
15. Cho YS, Moon HB. The role of oxidative stress in the pathogenesis of asthma. Allergy Asthma Immunol Res. 2010;2(3):183-7.
16. Moreno-Macias H, Romieu I. Effects of antioxidant supplements and nutrients on patients with asthma and allergies. J Allergy Clin Immun. 2014;133(5):1237-45.
17. Dauchet L, Amouyel P, Dallongeville J. Fruits, vegetables and coronary heart disease. Nat Rev Cardiol. 2009;6(9):599-608.
18. Seaton A, Godden DJ, Brown K. Increase in asthma: a more toxic environment or a more susceptible population? Thorax. 1994;49(2):171-4.
19. Devereux G, Seaton A. Diet as a risk factor for atopy and asthma. J Allergy Clin Immun. 2005;115(6):1109-17.
20. Ekstrom S, Magnusson J, Kull I, Lind T, Almqvist C, Melen E, et al. Maternal body mass index in early pregnancy and offspring asthma, rhinitis and eczema up to 16 years of age. Clin Exp Allergy. 2015;45(1):283-91.
21. Strachan DP. Hay fever, hygiene, and household size. BMJ. 1989;299(6710):1259-60.
22. Fuchs O, von Mutius E. Prenatal and childhood infections: implications for the development and treatment of childhood asthma. Lancet Respir Med. 2013;1(9):743-54.

23. Smits HH, van der Vlugt LE, von Mutius E, Hiemstra PS. Childhood allergies and asthma: New insights on environmental exposures and local immunity at the lung barrier. Curr Opin Immunol. 2016;42:41-7.
24. von Mutius E. The microbial environment and its influence on asthma prevention in early life. J Allergy Clin Immun. 2016;137(3):680-9.
25. Castro-Rodriguez JA, Forno E, Rodriguez-Martinez CE, Celedon JC. Risk and Protective Factors for Childhood Asthma: What Is the Evidence? J Allergy Clin Immunol Pract. 2016;4(6):1111-22.
26. Neuman A, Hohmann C, Orsini N, Pershagen G, Eller E, Kjaer HF, et al. Maternal smoking in pregnancy and asthma in preschool children: a pooled analysis of eight birth cohorts. Am J Respir Crit Care Med. 2012;186(10):1037-43.
27. Thacher JD, Gruzieva O, Pershagen G, Neuman A, Wickman M, Kull I, et al. Pre- and postnatal exposure to parental smoking and allergic disease through adolescence. Pediatrics. 2014;134(3):428-34.
28. Wright RJ, Brunst KJ. Programming of respiratory health in childhood: influence of outdoor air pollution. Curr Opin Pediatr. 2013;25(2):232-9.
29. Hsu HH, Chiu YH, Coull BA, Kloog I, Schwartz J, Lee A, et al. Prenatal Particulate Air Pollution and Asthma Onset in Urban Children. Identifying Sensitive Windows and Sex Differences. Am J Respir Crit Care Med. 2015;192(9):1052-9.
30. Keil T, Bockelbrink A, Reich A, Hoffmann U, Kamin W, Forster J, et al. The natural history of allergic rhinitis in childhood. Pediatr Allergy Immunol. 2010;21(6):962-9.
31. Ballardini N, Kull I, Lind T, Hallner E, Almqvist C, Ostblom E, et al. Development and comorbidity of eczema, asthma and rhinitis to age 12: data from the BAMSE birth cohort. Allergy. 2012;67(4):537-44.
32. Westman M, Stjarne P, Bergstrom A, Kull I, Toskala E, Cardell LO, et al. Chronic rhinosinusitis is rare but bothersome in adolescents from a Swedish population-based cohort. J Allergy Clin Immunol. 2015;136(2):512-4 e6.
33. Mudarri D, Fisk WJ. Public health and economic impact of dampness and mold. Indoor Air. 2007;17(3):226-35.
34. World Health Organization. WHO guidelines for indoor air quality: dampness and mould. Copenhagen: World Health Organization; 2009. [cited 2017 Mar 21]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0017/43325/E92645.pdf?ua=1.
35. Cho SJ, Cox-Ganser JM, Park JH. Observational scores of dampness and mold associated with measurements of microbial agents and moisture in three public schools. Indoor Air. 2016;26(2):168-78.
36. Mendell MJ, Mirer AG, Cheung K, Tong M, Douwes J. Respiratory and allergic health effects of dampness, mold, and dampness-related agents: a review of the epidemiologic evidence. Environ Health Perspect. 2011;119(6):748-56.
37. Fisk WJ, Lei-Gomez Q, Mendell MJ. Meta-analyses of the associations of respiratory health effects with dampness and mold in homes. Indoor Air. 2007;17(4):284-96.
38. Quansah R, Jaakkola MS, Hugg TT, Heikkinen SA, Jaakkola JJ. Residential dampness and molds and the risk of developing asthma: a systematic review and meta-analysis. PLoS One. 2012;7(11):e47526.
39. Jaakkola MS, Quansah R, Hugg TT, Heikkinen SA, Jaakkola JJ. Association of indoor dampness and molds with rhinitis risk: a systematic review and meta-analysis. J Allergy Clin Immunol. 2013;132(5):1099-110 e18.
40. Tischer C, Chen CM, Heinrich J. Association between domestic mould and mould components, and asthma and allergy in children: a systematic review. Eur Respir J. 2011;38(4):812-24.
41. Su MW, Tung KY, Liang PH, Tsai CH, Kuo NW, Lee YL. Gene-gene and gene-environmental interactions of childhood asthma: a multifactor dimension reduction approach. PLoS One. 2012;7(2):e30694.
42. World Health Organization. Ambient air pollution: A global assessment of exposure and burden of disease [internet]. Geneva: World Health Organization; 2016. [cited 2017 Mar 21]. Available from: http://apps.who.int/iris/bitstream/10665/250141/1/9789241511353-eng.pdf?ua=1.

43. Perez L, Rapp R, Kunzli N. The Year of the Lung: outdoor air pollution and lung health. Swiss Med Wkly. 2010;140:w13129.
44. World Health Organization. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment [internet]. Geneva: World Health Organization; 2006. [cited 2017 Mar 21]. Available from: http://apps.who.int/iris/bitstream/10665/69477/1/WHO_SDE_PHE_OEH_06.02_eng.pdf.
45. Esposito S, Tenconi R, Lelii M, Preti V, Nazzari E, Consolo S, et al. Possible molecular mechanisms linking air pollution and asthma in children. BMC Pulm Med. 2014;14:31.
46. Anderson JO, Thundiyil JG, Stolbach A. Clearing the air: a review of the effects of particulate matter air pollution on human health. J Med Toxicol. 2012;8(2):166-75.
47. European Commission. Air Quality Standards [internet]. Brussels: European Commission; 2017. [cited 2017 Mar 6]. Available from: http://ec.europa.eu/environment/air/quality/standards.htm.
48. Ordinance on ambient air quality (SFS 2010:477) [internet]. Stockholm: Ministry of Environment and Energy [cited 2017 Mar 21]. Available from: http://swedishepa.se/upload/in-english/legislation/the-swedish-environmental-code/environmental-quality-standards/air-quality-ordinance-sfs-2010-477.pdf.
49. Schultz ES, Gruzieva O, Bellander T, Bottai M, Hallberg J, Kull I, et al. Traffic-related air pollution and lung function in children at 8 years of age: a birth cohort study. Am J Respir Crit Care Med. 2012;186(12):1286-91.
50. Gauderman WJ, Avol E, Gilliland F, Vora H, Thomas D, Berhane K, et al. The effect of air pollution on lung development from 10 to 18 years of age. N Engl J Med. 2004;351(11):1057-67.
51. Gehring U, Gruzieva O, Agius RM, Beelen R, Custovic A, Cyrys J, et al. Air pollution exposure and lung function in children: the ESCAPE project. Environ Health Perspect. 2013;121(11-12):1357-64.
52. Bowatte G, Lodge C, Lowe AJ, Erbas B, Perret J, Abramson MJ, et al. The influence of childhood traffic-related air pollution exposure on asthma, allergy and sensitization: a systematic review and a meta-analysis of birth cohort studies. Allergy. 2015;70(3):245-56.
53. HEI. Panel on the Health Effects of Traffic-Related Air Pollution. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. HEI Special report 17. Health Effects Institute, Boston, MA; 2010.
54. Gehring U, Wijga AH, Hoek G, Bellander T, Berdel D, Bruske I, et al. Exposure to air pollution and development of asthma and rhinoconjunctivitis throughout childhood and adolescence: a population-based birth cohort study. Lancet Respir Med. 2015;3(12):933-42.
55. Khreis H, Kelly C, Tate J, Parslow R, Lucas K, Nieuwenhuijsen M. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. Environ Int. 2017;100:1-31.
56. Gruzieva O, Gehring U, Aalberse R, Agius R, Beelen R, Behrendt H, et al. Meta-analysis of air pollution exposure association with allergic sensitization in European birth cohorts. J Allergy Clin Immunol. 2014;133(3):767-76 e7.
57. Ministers NCo. Nordic Nutrition Recommendations 2012: Integrating Nutrition and Physical Activity. Nordic Council of Ministers; 2014.
58. Allan K, Kelly FJ, Devereux G. Antioxidants and allergic disease: a case of too little or too much? Clin Exp Allergy. 2010;40(3):370-80.
59. Comhair SA, Erzurum SC. Antioxidant responses to oxidant-mediated lung diseases. Am J Physiol Lung Cell Mol Physiol. 2002;283(2):L246-55.
60. Maslova E, Hansen S, Strom M, Halldorsson TI, Olsen SF. Maternal intake of vitamins A, E and K in pregnancy and child allergic disease: a longitudinal study from the Danish National Birth Cohort. Br J Nutr. 2014;111(6):1096-108.
61. Miller DR, Turner SW, Spiteri-Cornish D, Scaife AR, Danielian PJ, Devereux GS, et al. Maternal vitamin D and E intakes during early pregnancy are associated with airway epithelial cell responses in neonates. Clin Exp Allergy. 2015;45(5):920-7.
62. Netting MJ, Middleton PF, Makrides M. Does maternal diet during pregnancy and lactation affect outcomes in offspring? A systematic review of food-based approaches. Nutrition. 2014;30(11-12):1225-41.

63. West CE, Dunstan J, McCarthy S, Metcalfe J, D'Vaz N, Meldrum S, et al. Associations between maternal antioxidant intakes in pregnancy and infant allergic outcomes. Nutrients. 2012;4(11):1747-58.
64. Matheson MC, Allen KJ, Tang ML. Understanding the evidence for and against the role of breastfeeding in allergy prevention. Clin Exp Allergy. 2012;42(6):827-51.
65. Ierodiakonou D, Garcia-Larsen V, Logan A, Groome A, Cunha S, Chivinge J, et al. Timing of Allergenic Food Introduction to the Infant Diet and Risk of Allergic or Autoimmune Disease: A Systematic Review and Meta-analysis. JAMA. 2016;316(11):1181-92.
66. Nurmatov U, Devereux G, Sheikh A. Nutrients and foods for the primary prevention of asthma and allergy: systematic review and meta-analysis. J Allergy Clin Immunol. 2011;127(3):724-33 e1-30.
67. Garcia-Larsen V, Del Giacco SR, Moreira A, Bonini M, Charles D, Reeves T, et al. Asthma and dietary intake: an overview of systematic reviews. Allergy. 2016;71(4):433-42.
68. Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M, Moreno-Macias H, Reyes-Ruiz NI, Estela del Rio-Navarro B, et al. Genetic polymorphism of GSTM1 and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. Thorax. 2004;59(1):8-10.
69. Han YY, Forno E, Holguin F, Celedon JC. Diet and asthma: an update. Curr Opin Allergy Clin Immunol. 2015;15(4):369-74.
70. Enghart Barbieri H, Pearson M, Becker W. Riksmaten-Barn 2003. Livsmedels- och näringsintag bland barn i Sverige. Uppsala; 2003.
71. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014;36:e2014009.
72. Serafini M, Del Rio D. Understanding the association between dietary antioxidants, redox status and disease: is the Total Antioxidant Capacity the right tool? Redox Rep. 2004;9(3):145-52.
73. Rosenlund H, Kull I, Pershagen G, Wolk A, Wickman M, Bergstrom A. Fruit and vegetable consumption in relation to allergy: disease-related modification of consumption? J Allergy Clin Immunol. 2011;127(5):1219-25.
74. Lodovici M, Bigagli E. Oxidative stress and air pollution exposure. J Toxicol. 2011;2011:487074.
75. Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. Annu Rev Pharmacol Toxicol. 2005;45:51-88.
76. Fryer AA, Hume R, Strange RC. The development of glutathione S-transferase and glutathione peroxidase activities in human lung. Biochim Biophys Acta. 1986;883(3):448-53.
77. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. Genes Immun. 2006;7(2):95-100.
78. Minelli C, Granell R, Newson R, Rose-Zerilli MJ, Torrent M, Ring SM, et al. Glutathione-S-transferase genes and asthma phenotypes: a Human Genome Epidemiology (HuGE) systematic review and meta-analysis including unpublished data. Int J Epidemiol. 2010;39(2):539-62.
79. Romieu I, Ramirez-Aguilar M, Sienra-Monge JJ, Moreno-Macias H, del Rio-Navarro BE, David G, et al. GSTM1 and GSTP1 and respiratory health in asthmatic children exposed to ozone. Eur Respir J. 2006;28(5):953-9.
80. Salam MT, Lin PC, Avol EL, Gauderman WJ, Gilliland FD. Microsomal epoxide hydrolase, glutathione S-transferase P1, traffic and childhood asthma. Thorax. 2007;62(12):1050-7.
81. Melen E, Nyberg F, Lindgren CM, Berglind N, Zucchelli M, Nordling E, et al. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. Environ Health Perspect. 2008;116(8):1077-84.
82. Schroer KT, Myers JMB, Ryan PH, LeMasters GK, Bernstein DI, Villareal M, et al. Associations between Multiple Environmental Exposures and Glutathione S-transferase P1 on Persistent Wheezing in a Birth Cohort. J Pediatr-Us. 2009;154(3):401-8.
83. Hwang BF, Young LH, Tsai CH, Tung KY, Wang PC, Su MW, et al. Fine particle, ozone exposure, and asthma/wheezing: effect modification by glutathione S-transferase P1 polymorphisms. PLoS One. 2013;8(1):e52715.

84. MacIntyre EA, Brauer M, Melen E, Bauer CP, Bauer M, Berdel D, et al. GSTP1 and TNF Gene variants and associations between air pollution and incident childhood asthma: the traffic, asthma and genetics (TAG) study. Environ Health Perspect. 2014;122(4):418-24.
85. Reddy P, Naidoo RN, Robins TG, Mentz G, Li H, London SJ, et al. GSTM1 and GSTP1 gene variants and the effect of air pollutants on lung function measures in South African children. Am J Ind Med. 2012;55(12):1078-86.
86. Aoki T, Hirota T, Tamari M, Ichikawa K, Takeda K, Arinami T, et al. An association between asthma and TNF-308G/A polymorphism: meta-analysis. J Hum Genet. 2006;51(8):677-85.
87. Wu H, Romieu I, Sienra-Monge JJ, del Rio-Navarro BE, Anderson DM, Dunn EW, et al. Parental smoking modifies the relation between genetic variation in tumor necrosis factor-alpha (TNF) and childhood asthma. Environ Health Perspect. 2007;115(4):616-22.
88. Panasevich S, Lindgren C, Kere J, Wickman M, Pershagen G, Nyberg F, et al. Interaction between early maternal smoking and variants in TNF and GSTP1 in childhood wheezing. Clin Exp Allergy. 2010;40(3):458-67.
89. Xu YL, Reinscheid RK, Huitron-Resendiz S, Clark SD, Wang Z, Lin SH, et al. Neuropeptide S: a neuropeptide promoting arousal and anxiolytic-like effects. Neuron. 2004;43(4):487-97.
90. Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, et al. Characterization of a common susceptibility locus for asthma-related traits. Science. 2004;304(5668):300-4.
91. Melen E, Bruce S, Doekes G, Kabesch M, Laitinen T, Lauener R, et al. Haplotypes of G protein-coupled receptor 154 are associated with childhood allergy and asthma. Am J Respir Crit Care Med. 2005;171(10):1089-95.
92. Kormann MS, Carr D, Klopp N, Illig T, Leupold W, Fritzsch C, et al. G-Protein-coupled receptor polymorphisms are associated with asthma in a large German population. Am J Respir Crit Care Med. 2005;171(12):1358-62.
93. Feng Y, Hong X, Wang L, Jiang S, Chen C, Wang B, et al. G protein-coupled receptor 154 gene polymorphism is associated with airway hyperresponsiveness to methacholine in a Chinese population. J Allergy Clin Immunol. 2006;117(3):612-7.
94. Malerba G, Lindgren CM, Xumerle L, Kiviluoma P, Trabetti E, Laitinen T, et al. Chromosome 7p linkage and GPR154 gene association in Italian families with allergic asthma. Clin Exp Allergy. 2007;37(1):83-9.
95. Hersh CP, Raby BA, Soto-Quiros ME, Murphy AJ, Avila L, Lasky-Su J, et al. Comprehensive testing of positionally cloned asthma genes in two populations. Am J Respir Crit Care Med. 2007;176(9):849-57.
96. Daley D, Lemire M, Akhabir L, Chan-Yeung M, He JQ, McDonald T, et al. Analyses of associations with asthma in four asthma population samples from Canada and Australia. Hum Genet. 2009;125(4):445-59.
97. Vergara C, Jimenez S, Acevedo N, Martinez B, Mercado D, Gusmao L, et al. Association of G-protein-coupled receptor 154 with asthma and total IgE in a population of the Caribbean coast of Colombia. Clin Exp Allergy. 2009;39(10):1558-68.
98. Castro-Giner F, de Cid R, Gonzalez JR, Jarvis D, Heinrich J, Janson C, et al. Positionally cloned genes and age-specific effects in asthma and atopy: an international population-based cohort study (ECRHS). Thorax. 2010;65(2):124-31.
99. Wu H, Romieu I, Sienra-Monge JJ, del Rio-Navarro BE, Burdett L, Yuenger J, et al. Lack of association between genetic variation in G-protein-coupled receptor for asthma susceptibility and childhood asthma and atopy. Genes Immun. 2008;9(3):224-30.
100. Blakey JD, Sayers I, Ring SM, Strachan DP, Hall IP. Positionally cloned asthma susceptibility gene polymorphisms and disease risk in the British 1958 Birth Cohort. Thorax. 2009;64(5):381-7.
101. Bruce S, Nyberg F, Melen E, James A, Pulkkinen V, Orsmark-Pietras C, et al. The protective effect of farm animal exposure on childhood allergy is modified by NPSR1 polymorphisms. J Med Genet. 2009;46(3):159-67.
102. International HapMap C, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007;449(7164):851-61.
103. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.

104. Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.
105. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature. 2007;448(7152):470-3.
106. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med. 2010;363(13):1211-21.
107. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev. 2011;242(1):10-30.
108. Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. N Engl J Med. 2008;359(19):1985-94.
109. Caliskan M, Bochkov YA, Kreiner-Moller E, Bonnelykke K, Stein MM, Du G, et al. Rhinovirus wheezing illness and genetic risk of childhood-onset asthma. N Engl J Med. 2013;368(15):1398-407.
110. Vercelli D. Gene-environment interactions in asthma and allergy: the end of the beginning? Curr Opin Allergy Clin Immunol. 2010;10(2):145-8.
111. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP. Finding novel genes by testing G x E interactions in a genome-wide association study. Genet Epidemiol. 2013;37(6):603-13.
112. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. Hum Hered. 2007;63(2):111-9.
113. Murcray CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. Am J Epidemiol. 2009;169(2):219-26.
114. Fuertes E, Brauer M, MacIntyre E, Bauer M, Bellander T, von Berg A, et al. Childhood allergic rhinitis, traffic-related air pollution, and variability in the GSTP1, TNF, TLR2, and TLR4 genes: results from the TAG Study. J Allergy Clin Immunol. 2013;132(2):342-52 e2.
115. Schubeler D. Function and information content of DNA methylation. Nature. 2015;517(7534):321-6.
116. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011;25(10):1010-22.
117. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One. 2012;7(7):e41361.
118. Non AL, Thayer ZM. Epigenetics for anthropologists: An introduction to methods. Am J Hum Biol. 2015;27(3):295-303.
119. Lockett GA, Patil VK, Soto-Ramirez N, Ziyab AH, Holloway JW, Karmaus W. Epigenomics and allergic disease. Epigenomics. 2013;5(6):685-99.
120. Gruzieva O, Merid SK, Melen E. An update on epigenetics and childhood respiratory diseases. Paediatr Respir Rev. 2014;15(4):348-54.
121. Devries A, Vercelli D. Epigenetics of human asthma and allergy: promises to keep. Asian Pac J Allergy Immunol. 2013;31(3):183-9.
122. Harb H, Alashkar Alhamwe B, Garn H, Renz H, Potaczek DP. Recent developments in epigenetics of pediatric asthma. Curr Opin Pediatr. 2016;28(6):754-63.
123. Gruzieva O, Xu CJ, Breton CV, Annesi-Maesano I, Anto JM, Auffray C, et al. Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO2 Air Pollution Exposure. Environ Health Perspect. 2017;125(1):104-10.
124. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet. 2016;98(4):680-96.
125. Xu CJ, Bonder MJ, Soderhall C, Bustamante M, Baiz N, Gehring U, et al. The emerging landscape of dynamic DNA methylation in early childhood. BMC Genomics. 2017;18(1):25.
126. Kirkwood BR, Sterne AC. Essential Medical Statistics. 2nd ed: Blackwell Science Ltd, Oxford, UK; 2003.
127. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005;95 Suppl 1:S144-50.
128. Rothman KJ. Epidemiology: an introduction. New York, USA: Oxford university press Inc; 2002.

129. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: presentation of a prospective longitudinal birth cohort study. Pediatr Allergy Immunol. 2002;13 Suppl 15:11-3.
130. MacIntyre EA, Carlsten C, MacNutt M, Fuertes E, Melen E, Tiesler CM, et al. Traffic, asthma and genetics: combining international birth cohort data to examine genetics as a mediator of traffic-related air pollution's impact on childhood asthma. Eur J Epidemiol. 2013;28(7):597-606.
131. Berg A, Kramer U, Link E, Bollrath C, Heinrich J, Brockow I, et al. Impact of early feeding on childhood eczema: development after nutritional intervention compared with the natural course - the GINIplus study up to the age of 6 years. Clin Exp Allergy. 2010;40(4):627-36.
132. Heinrich J, Bolte G, Holscher B, Douwes J, Lehmann I, Fahlbusch B, et al. Allergens and endotoxin on mothers' mattresses and total immunoglobulin E in cord blood of neonates. European Respiratory Journal. 2002;20(3):617-23.
133. Wijga AH, Kerkhof M, Gehring U, de Jongste JC, Postma DS, Aalberse RC, et al. Cohort profile: the prevention and incidence of asthma and mite allergy (PIAMA) birth cohort. Int J Epidemiol. 2014;43(2):527-35.
134. Carlsten C, Dybuncio A, Becker A, Chan-Yeung M, Brauer M. Traffic-related air pollution and incident asthma in a high-risk birth cohort. Occup Environ Med. 2011;68(4):291-5.
135. Kozyrskyj AL, HayGlass KT, Sandford AJ, Pare PD, Chan-Yeung M, Becker AB. A novel study design to investigate the early-life origins of asthma in children (SAGE study). Allergy. 2009;64(8):1185-93.
136. Baker D, Henderson J. Differences between infants and adults in the social aetiology of wheeze. The ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. J Epidemiol Community Health. 1999;53(10):636-42.
137. Kupczyk M, Wenzel S. U.S. and European severe asthma cohorts: what can they teach us about severe asthma? J Intern Med. 2012;272(2):121-32.
138. Peters JM, Avol E, Gauderman WJ, Linn WS, Navidi W, London SJ, et al. A study of twelve Southern California communities with differing levels and types of air pollution. II. Effects on pulmonary function. Am J Respir Crit Care Med. 1999;159(3):768-75.
139. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet. 2012;8(11):e1003029.
140. Lamontagne M, Timens W, Hao K, Bosse Y, Laviolette M, Steiling K, et al. Genetic regulation of gene expression in the lung identifies CST3 and CD22 as potential causal genes for airflow obstruction. Thorax. 2014;69(11):997-1004.
141. Jiang R, Jones MJ, Sava F, Kobor MS, Carlsten C. Short-term diesel exhaust inhalation in a controlled human crossover study is associated with changes in DNA methylation of circulating mononuclear cells in asthmatics. Part Fibre Toxicol. 2014;11:71.
142. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45(6):580-5.
143. Severin J, Lizio M, Harshbarger J, Kawaji H, Daub CO, Hayashizaki Y, et al. Interactive visualization and analysis of large-scale sequencing datasets using ZENBU. Nat Biotechnol. 2014;32(3):217-9.
144. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. 2015;16:22.
145. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol. 2010;28(12):1248-50.
146. Nordling E, Berglind N, Melen E, Emenius G, Hallberg J, Nyberg F, et al. Traffic-related air pollution and childhood respiratory symptoms, function and allergies. Epidemiology. 2008;19(3):401-8.
147. Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ Health Persp. 2001;109(6):633-9.
148. Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. Atmos Environ. 2013;72:10-23.
149. Johansson C, Hadenius A, Johansson P-Å, Jonson T. SHAPE The Stockholm Study on Health Effects of Air Pollution and their Economic Consequences. Part I: NO2 and Particulate Matter in Stockholm- Concentrations and Pollution Exposure. Stockholm: Stockholm Environment

and Health Protection Administration, Air quality and Noise analysis; 1999. [cited 2017 Feb 14]. Available from: http://slb.nu/slb/rapporter/pdf6/lvf1999_Ex1.pdf.
150. Johansson C, Burman I. The Stockholm trial-Effects on Air quality and Health [internet]. Stockholm: SLB-analys at City of Stockholm Environment and Health Administration; 2006. [cited 2017 Feb 14]. Available from: http://www.stockholmsforsoket.se/upload/Sammanfattningar/English/Effects%20on%20air%20quality%20and%20health.pdf.
151. Hammarström U, Karlsson B. Costs and exhaust emissions for road planning (Fordonskostnader och avgasemissioner för vägplanering (EVA)). Linköping: Swedish National Road and Transport research Institute; 1994.
152. SMHI. Airviro [internet]. Norrköping: SMHI; 2017. [cited 2017 Mar 6]. Available from: http://www.smhi.se/airviro.
153. Allen RW, Amram O, Wheeler AJ, Brauer M. The transferability of NO and NO2 land use regression models between cities and pollutants. Atmos Environ. 2011;45(2):369-78.
154. Peters JM, Avol E, Navidi W, London SJ, Gauderman WJ, Lurmann F, et al. A study of twelve Southern California communities with differing levels and types of air pollution. I. Prevalence of respiratory morbidity. Am J Respir Crit Care Med. 1999;159(3):760-7.
155. McConnell R, Berhane K, Gilliland F, Molitor J, Thomas D, Lurmann F, et al. Prospective study of air pollution and bronchitic symptoms in children with asthma. Am J Respir Crit Care Med. 2003;168(7):790-7.
156. Reinius LE, Gref A, Saaf A, Acevedo N, Joerink M, Kupczyk M, et al. DNA methylation in the Neuropeptide S Receptor 1 (NPSR1) promoter in relation to asthma and environmental factors. PLoS One. 2013;8(1):e53877.
157. Slama R, Grabsch C, Lepeule J, Siroux V, Cyrys J, Sausenthaler S, et al. Maternal fine particulate matter exposure, polymorphism in xenobiotic-metabolizing genes and offspring birth weight. Reprod Toxicol. 2010;30(4):600-12.
158. Koopman LP, van Strien RT, Kerkhof M, Wijga A, Smit HA, de Jongste JC, et al. Placebo-controlled trial of house dust mite-impermeable mattress covers: effect on symptoms in early childhood. Am J Respir Crit Care Med. 2002;166(3):307-13.
159. Shaheen SO, Newson RB, Ring SM, Rose-Zerilli MJ, Holloway JW, Henderson AJ. Prenatal and infant acetaminophen exposure, antioxidant gene polymorphisms, and childhood asthma. J Allergy Clin Immunol. 2010;126(6):1141-8 e7.
160. Paternoster L, Standl M, Chen CM, Ramasamy A, Bonnelykke K, Duijts L, et al. Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. Nature Genetics. 2012;44(2):187-92.
161. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet. 2011;43(9):887-92.
162. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci U S A. 2005;102(44):15785-90.
163. Bergström L, Kylberg E, Hagman U, Erikson H, Bruce Å. The food composition database KOST: the National Food Administration's information system for nutritive valuse of food. Vår föda. 1991;43:439-47.
164. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
165. Wu X, Beecher GR, Holden JM, Haytowitz DB, Gebhardt SE, Prior RL. Lipophilic and hydrophilic antioxidant capacities of common foods in the United States. J Agric Food Chem. 2004;52(12):4026-37.
166. Prior RL, Wu X, Schaich K. Standardized methods for the determination of antioxidant capacity and phenolics in foods and dietary supplements. J Agric Food Chem. 2005;53(10):4290-302.
167. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.
168. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 2001;29(4):1165-88.

169. Hedges LV, Olkin I. The inverse chi-square method. Statistical methods for meta-analysis. San Diago, CA: Academic Press, Inc; 1985. p. 37-9.
170. Jaakkola JJ, Hwang BF, Jaakkola MS. Home dampness and molds as determinants of allergic rhinitis in childhood: a 6-year, population-based cohort study. Am J Epidemiol. 2010;172(4):451-9.
171. Karvonen AM, Hyvarinen A, Korppi M, Haverinen-Shaughnessy U, Renz H, Pfefferle PI, et al. Moisture damage and asthma: a birth cohort study. Pediatrics. 2015;135(3):e598-606.
172. Thacher JD, Gruzieva O, Pershagen G, Melen E, Lorentzen JC, Kull I, et al. Mold and dampness exposure and allergic outcomes from birth to adolescence: data from the BAMSE cohort. Allergy. 2016.
173. Tischer CG, Hohmann C, Thiering E, Herbarth O, Muller A, Henderson J, et al. Meta-analysis of mould and dampness exposure on asthma and allergy in eight European birth cohorts: an ENRIECO initiative. Allergy. 2011;66(12):1570-9.
174. Kim HJ, Lee E, Lee SH, Kang MJ, Hong SJ. Mold elicits atopic dermatitis by reactive oxygen species: Epidemiology and mechanism studies. Clin Immunol. 2015;161(2):384-90.
175. Hwang BF, Liu IP, Huang TP. Gene-environment interaction between interleukin-4 promoter and molds in childhood asthma. Ann Epidemiol. 2012;22(4):250-6.
176. Kim WK, Kwon JW, Seo JH, Kim HY, Yu J, Kim BJ, et al. Interaction between IL13 genotype and environmental factors in the risk for allergic rhinitis in Korean children. J Allergy Clin Immunol. 2012;130(2):421-6 e5.
177. Tsai CH, Tung KY, Chen CH, Lee YL. Tumour necrosis factor G-308A polymorphism modifies the effect of home dampness on childhood asthma. Occup Environ Med. 2011;68(10):771-6.
178. Wu AC, Lasky-Su J, Rogers CA, Klanderman BJ, Litonjua AA. Fungal exposure modulates the effect of polymorphisms of chitinases on emergency department visits and hospitalizations. Am J Respir Crit Care Med. 2010;182(7):884-9.
179. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet. 2010;42(1):45-52.
180. Panasevich S, Melen E, Hallberg J, Bergstrom A, Svartengren M, Pershagen G, et al. Investigation of novel genes for lung function in children and their interaction with tobacco smoke exposure: a preliminary report. Acta Paediatr. 2013;102(5):498-503.
181. Castaldi PJ, Cho MH, Litonjua AA, Bakke P, Gulsvik A, Lomas DA, et al. The association of genome-wide significant spirometric loci with chronic obstructive pulmonary disease susceptibility. Am J Respir Cell Mol Biol. 2011;45(6):1147-53.
182. Chen W, Brehm JM, Manichaikul A, Cho MH, Boutaoui N, Yan Q, et al. A genome-wide association study of chronic obstructive pulmonary disease in Hispanics. Ann Am Thorac Soc. 2015;12(3):340-8.
183. Martinez FD. Early-Life Origins of Chronic Obstructive Pulmonary Disease. N Engl J Med. 2016;375(9):871-8.
184. Scholtens S, Postma DS, Moffatt MF, Panasevich S, Granell R, Henderson AJ, et al. Novel childhood asthma genes interact with in utero and early-life tobacco smoke exposure. J Allergy Clin Immunol. 2014;133(3):885-8.
185. Ege MJ, Strachan DP, Cookson WO, Moffatt MF, Gut I, Lathrop M, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. J Allergy Clin Immunol. 2011;127(1):138-44, 44 e1-4.
186. Hancock DB, Soler Artigas M, Gharib SA, Henry A, Manichaikul A, Ramasamy A, et al. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. PLoS Genet. 2012;8(12):e1003098.
187. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Soler Artigas M, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. Lancet Respir Med. 2015;3(10):769-81.
188. Schieck M, Schouten JP, Michel S, Suttner K, Toncheva AA, Gaertner VD, et al. Doublesex and mab-3 related transcription factor 1 (DMRT1) is a sex-specific genetic determinant of childhood-onset asthma and is expressed in testis and macrophages. J Allergy Clin Immunol. 2016;138(2):421-31.

189. Nieuwenhuis MA, Siedlinski M, van den Berge M, Granell R, Li X, Niens M, et al. Combining genomewide association study and lung eQTL analysis provides evidence for novel genes associated with asthma. Allergy. 2016;71(12):1712-20.
190. de Jong K, Vonk JM, Timens W, Bosse Y, Sin DD, Hao K, et al. Genome-wide interaction study of gene-by-occupational exposure and effects on FEV1 levels. J Allergy Clin Immunol. 2015;136(6):1664-72 e1-14.
191. Imboden M, Kumar A, Curjuric I, Adam M, Thun GA, Haun M, et al. Modification of the association between PM10 and lung function decline by cadherin 13 polymorphisms in the SAPALDIA cohort: a genome-wide interaction analysis. Environ Health Perspect. 2015;123(1):72-9.
192. Forno E, Sordillo J, Brehm J, Chen W, Benos T, Yan Q, et al. Genome-wide interaction study of dust mite allergen on lung function in children with asthma. J Allergy Clin Immunol. 2017.
193. Saadeh D, Salameh P, Baldi I, Raherison C. Diet and allergic diseases among population aged 0 to 18 years: myth or reality? Nutrients. 2013;5(9):3399-423.
194. Rosenlund H, Magnusson J, Kull I, Hakansson N, Wolk A, Pershagen G, et al. Antioxidant intake and allergic disease in children. Clin Exp Allergy. 2012;42(10):1491-500.
195. Nakamura K, Wada K, Sahashi Y, Tamai Y, Tsuji M, Watanabe K, et al. Associations of intake of antioxidant vitamins and fatty acids with asthma in pre-school children. Public Health Nutr. 2013;16(11):2040-5.
196. Seo JH, Kwon SO, Lee SY, Kim HY, Kwon JW, Kim BJ, et al. Association of antioxidants with allergic rhinitis in children from seoul. Allergy Asthma Immunol Res. 2013;5(2):81-7.
197. Papadopoulou A, Panagiotakos DB, Hatziagorou E, Antonogeorgos G, Matziou VN, Tsanakas JN, et al. Antioxidant foods consumption and childhood asthma and other allergic diseases: The Greek cohorts of the ISAAC II survey. Allergol Immunopathol (Madr). 2015;43(4):353-60.
198. Patel S, Murray CS, Woodcock A, Simpson A, Custovic A. Dietary antioxidant intake, allergic sensitization and allergic diseases in young children. Allergy. 2009;64(12):1766-72.
199. Hamalainen N, Nwaru BI, Erlund I, Takkinen HM, Ahonen S, Toppari J, et al. Serum carotenoid and tocopherol concentrations and risk of asthma in childhood: a nested case-control study. Clin Exp Allergy. 2017;47(3):401-9.
200. Rodriguez-Rodriguez E, Ortega RM, Gonzalez-Rodriguez LG, Penas-Ruiz C, Rodriguez-Rodriguez P. Dietary total antioxidant capacity and current asthma in Spanish schoolchildren: a case control-control study. Eur J Pediatr. 2014;173(4):517-23.
201. di Giuseppe R, Arcari A, Serafini M, Di Castelnuovo A, Zito F, De Curtis A, et al. Total dietary antioxidant capacity and lung function in an Italian population: a favorable role in premenopausal/never smoker women. Eur J Clin Nutr. 2012;66(1):61-8.
202. Romieu I, Barraza-Villarreal A, Escamilla-Nunez C, Texcalac-Sangrador JL, Hernandez-Cadena L, Diaz-Sanchez D, et al. Dietary intake, lung function and airway inflammation in Mexico City school children exposed to air pollutants. Respir Res. 2009;10:122.
203. Lee SY, Kim BS, Kwon SO, Oh SY, Shin HL, Jung YH, et al. Modification of additive effect between vitamins and ETS on childhood asthma risk according to GSTP1 polymorphism: a cross -sectional study. BMC Pulm Med. 2015;15:125.
204. Moreno-Macias H, Dockery DW, Schwartz J, Gold DR, Laird NM, Sienra-Monge JJ, et al. Ozone exposure, vitamin C intake, and genetic susceptibility of asthmatic children in Mexico City: a cohort study. Respir Res. 2013;14:14.
205. Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T, et al. A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. Nat Genet. 2001;28(1):87-91.
206. Andiappan AK, Wang de Y, Anantharaman R, Parate PN, Suri BK, Low HQ, et al. Genome-wide association study for atopy and allergic rhinitis in a Singapore Chinese population. PLoS One. 2011;6(5):e19719.
207. Chang JC, Kuo HC, Hsu TY, Ou CY, Liu CA, Chuang H, et al. Different genetic associations of the IgE production among fetus, infancy and childhood. PLoS One. 2013;8(8):e70362.
208. Everson TM, Lyons G, Zhang H, Soto-Ramirez N, Lockett GA, Patil VK, et al. DNA methylation loci associated with atopy and high serum IgE: a genome-wide application of recursive Random Forest feature selection. Genome Med. 2015;7:89.

209. Liang L, Willis-Owen SA, Laprise C, Wong KC, Davies GA, Hudson TJ, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. Nature. 2015;520(7549):670-4.
210. Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, et al. DNA methylation and childhood asthma in the inner city. J Allergy Clin Immunol. 2015;136(1):69-80.
211. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. Clin Epigenetics. 2015;7:113.
212. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med. 2009;180(5):462-7.
213. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ Health Perspect. 2012;120(10):1425-31.
214. Breton CV, Siegmund KD, Joubert BR, Wang X, Qui W, Carey V, et al. Prenatal tobacco smoke exposure is associated with childhood DNA CpG methylation. PLoS One. 2014;9(6):e99716.
215. Miller RL, Peden DB. Environmental effects on immune responses in patients with atopy and asthma. J Allergy Clin Immunol. 2014;134(5):1001-8.
216. Fu A, Leaderer BP, Gent JF, Leaderer D, Zhu Y. An environmental epigenetic study of ADRB2 5'-UTR methylation and childhood asthma severity. Clin Exp Allergy. 2012;42(11):1575-81.
217. Hew KM, Walker AI, Kohli A, Garcia M, Syed A, McDonald-Hyman C, et al. Childhood exposure to ambient polycyclic aromatic hydrocarbons is linked to epigenetic modifications and impaired systemic immunity in T cells. Clin Exp Allergy. 2015;45(1):238-48.
218. Nadeau K, McDonald-Hyman C, Noth EM, Pratt B, Hammond SK, Balmes J, et al. Ambient air pollution impairs regulatory T-cell function in asthma. J Allergy Clin Immunol. 2010;126(4):845-52 e10.
219. Somineni HK, Zhang X, Biagini Myers JM, Kovacic MB, Ulm A, Jurcak N, et al. Ten-eleven translocation 1 (TET1) methylation is associated with childhood asthma and traffic-related air pollution. J Allergy Clin Immunol. 2016;137(3):797-805 e5.
220. Brunst KJ, Leung YK, Ryan PH, Khurana Hershey GK, Levin L, Ji H, et al. Forkhead box protein 3 (FOXP3) hypermethylation is associated with diesel exhaust exposure and risk for childhood asthma. J Allergy Clin Immunol. 2013;131(2):592-4 e1-3.
221. Breton CV, Salam MT, Wang X, Byun HM, Siegmund KD, Gilliland FD. Particulate matter, DNA methylation in nitric oxide synthase, and childhood respiratory disease. Environ Health Perspect. 2012;120(9):1320-6.
222. Salam MT, Byun HM, Lurmann F, Breton CV, Wang X, Eckel SP, et al. Genetic and epigenetic variations in inducible nitric oxide synthase promoter, particulate pollution, and exhaled nitric oxide levels in children. J Allergy Clin Immunol. 2012;129(1):232-9 e1-7.
223. Perera F, Tang WY, Herbstman J, Tang D, Levin L, Miller R, et al. Relation of DNA methylation of 5'-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. PLoS One. 2009;4(2):e4488.
224. Tang WY, Levin L, Talaska G, Cheung YY, Herbstman J, Tang D, et al. Maternal exposure to polycyclic aromatic hydrocarbons and 5'-CpG methylation of interferon-gamma in cord white blood cells. Environ Health Perspect. 2012;120(8):1195-200.
225. Janssen BG, Godderis L, Pieters N, Poels K, Kicinski M, Cuypers A, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. Part Fibre Toxicol. 2013;10:22.
226. Liu J, Ballaney M, Al-alem U, Quan C, Jin X, Perera F, et al. Combined inhaled diesel exhaust particles and allergen exposure alter methylation of T helper genes and IgE production in vivo. Toxicol Sci. 2008;102(1):76-81.
227. Rossnerova A, Tulupova E, Tabashidze N, Schmuczerova J, Dostal M, Rossner P, Jr., et al. Factors affecting the 27K DNA methylation pattern in asthmatic and healthy children from locations with various environments. Mutat Res. 2013;741-742:18-26.
228. Ricceri F, Trevisan M, Fiano V, Grasso C, Fasanelli F, Scoccianti C, et al. Seasonality modifies methylation profiles in healthy people. PLoS One. 2014;9(9):e106846.

229. Huang YC, Karoly ED, Dailey LA, Schmitt MT, Silbajoris R, Graff DW, et al. Comparison of gene expression profiles induced by coarse, fine, and ultrafine particulate matter. J Toxicol Environ Health A. 2011;74(5):296-312.
230. van Leeuwen DM, van Herwijnen MH, Pedersen M, Knudsen LE, Kirsch-Volders M, Sram RJ, et al. Genome-wide differential gene expression in children exposed to air pollution in the Czech Republic. Mutat Res. 2006;600(1-2):12-22.
231. Huang YC, Schmitt M, Yang Z, Que LG, Stewart JC, Frampton MW, et al. Gene expression profile in circulating mononuclear cells after exposure to ultrafine carbon particles. Inhal Toxicol. 2010;22(10):835-46.
232. Peretz A, Peck EC, Bammler TK, Beyer RP, Sullivan JH, Trenga CA, et al. Diesel exhaust inhalation and assessment of peripheral blood mononuclear cell gene transcription effects: an exploratory study of healthy human volunteers. Inhal Toxicol. 2007;19(14):1107-19.
233. Rider CF, Yamamoto M, Gunther OP, Hirota JA, Singh A, Tebbutt SJ, et al. Controlled diesel exhaust and allergen coexposure modulates microRNA and gene expression in humans: Effects on inflammatory lung markers. J Allergy Clin Immunol. 2016;138(6):1690-700.
234. Pettit AP, Brooks A, Laumbach R, Fiedler N, Wang Q, Strickland PO, et al. Alteration of peripheral blood monocyte gene expression in humans following diesel exhaust inhalation. Inhal Toxicol. 2012;24(3):172-81.
235. Tischer CG, Heinrich J. Exposure assessment of residential mould, fungi and microbial components in relation to children's health: achievements and challenges. Int J Hyg Environ Health. 2013;216(2):109-14.
236. Emenius G, Svartengren M, Korsgaard J, Nordvall L, Pershagen G, Wickman M. Indoor exposures and recurrent wheezing in infants: a study in the BAMSE cohort. Acta Paediatr. 2004;93(7):899-905.
237. Eneroth K, Johansson C, Bellander T. Exposure - Comparison between Measurements and Calculations Based on Dispersion Modelling (Expose). Stockholm: SLB-analys; 2006. [cited 2017 Feb 9]. Available from: http://www.slb.nu/slb/rapporter/pdf/lvf2006_12.pdf.
238. Johansson C, Andersson C, Bergström R, Krecl P. Exposure to particles due to local and non-local sources in Stockholm - Estimates based on modelling and measurements 1997 - 2006 [internet]. Stockholm: Department of Applied Environmental Science, Stockholm University; 2008. [cited 2017 feb 10]. Available from: http://www.slb.nu/slb/rapporter/pdf8/itm2008_175.pdf.
239. Cesaroni G, Porta D, Badaloni C, Stafoggia M, Eeftens M, Meliefste K, et al. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. Environ Health. 2012;11:48.
240. Wang RR, Henderson SB, Sbihi H, Allen RW, Brauer M. Temporal stability of land use regression models for traffic-related air pollution. Atmos Environ. 2013;64:312-9.
241. Bellander T, Wichmann J, Lind T. Individual exposure to NO2 in relation to spatial and temporal exposure indices in Stockholm, Sweden: the INDEX study. PLoS One. 2012;7(6):e39536.
242. de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ Int. 2014;73:382-92.
243. Rautiainen S, Serafini M, Morgenstern R, Prior RL, Wolk A. The validity and reproducibility of food-frequency questionnaire-based total antioxidant capacity estimates in Swedish women. Am J Clin Nutr. 2008;87(5):1247-53.
244. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. Eur Respir J. 1995;8(3):483-91.
245. Okubo H, Syddall HE, Phillips DI, Sayer AA, Dennison EM, Cooper C, et al. Dietary total antioxidant capacity is related to glucose tolerance in older people: the Hertfordshire Cohort Study. Nutr Metab Cardiovasc Dis. 2014;24(3):301-8.
246. Burkett K, Greenwood C. A sequence of methodological changes due to sequencing. Curr Opin Allergy Clin Immunol. 2013;13(5):470-7.
247. Genereux DP, Johnson WC, Burden AF, Stoger R, Laird CD. Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. Nucleic Acids Res. 2008;36(22):e150.

248. Zhang H, Wang F, Kranzler HR, Yang C, Xu H, Wang Z, et al. Identification of methylation quantitative trait loci (mQTLs) influencing promoter DNA methylation of alcohol dependence risk genes. Hum Genet. 2014;133(9):1093-104.
249. Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin TE, Chatterjee N, et al. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Hum Genet. 2012;131(10):1591-613.
250. Ege MJ, Strachan DP. Comparisons of power of statistical methods for gene-environment interaction analyses. Eur J Epidemiol. 2013;28(10):785-97.
251. Kauffmann F, Demenais F. Gene-environment interactions in asthma and allergic diseases: challenges and perspectives. J Allergy Clin Immunol. 2012;130(6):1229-40; quiz 41-2.
252. Thomas DC, Lewinger JP, Murcray CE, Gauderman WJ. Invited commentary: GE-Whiz! Ratcheting gene-environment studies up to the whole genome and the whole exposome. Am J Epidemiol. 2012;175(3):203-7; discussion 8-9.
253. van der Sijde MR, Ng A, Fu J. Systems genetics: From GWAS to disease pathways. Biochim Biophys Acta. 2014;1842(10):1903-9.
254. Gibney ER, Nolan CM. Epigenetics and gene expression. Heredity (Edinb). 2010;105(1):4-13.
255. Su MW, Tsai CH, Tung KY, Hwang BF, Liang PH, Chiang BL, et al. GSTP1 is a hub gene for gene-air pollution interactions on childhood asthma. Allergy. 2013;68(12):1614-7.